

# DOCUMENT RESUME

TM 000 752

ED 052 258

AUTHOR  
TITLE  
INSTITUTION  
PUB DATE  
NOTE

Loye, David, Ed.  
Research at ETS: Projects and Publications.  
Educational Testing Service, Princeton, N.J.  
70  
114p.

EDRS PRICE  
DESCRIPTORS

EDRS Price MF-\$0.65 HC-\$6.58  
Careers, Cognitive Processes, Computer Programs,  
Disadvantaged Youth, Early Childhood Education,  
\*Educational Research, Graduate Study, Higher  
Education, Human Development, Individual  
Differences, Infants, Learning Theories, \*Literature  
Reviews, \*Measurement, Paired Associate Learning,  
Personality, Professional Education, \*Psychological  
Studies, Psychometrics, Research Methodology,  
Secondary Education, Social Behavior, \*Testing,  
Tests, Verbal Learning, Vocational Education

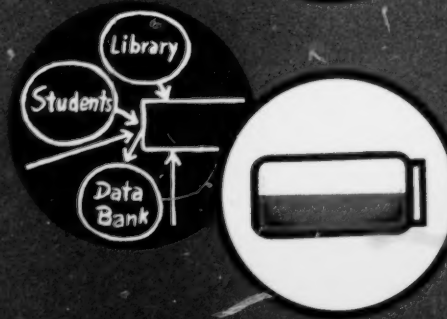
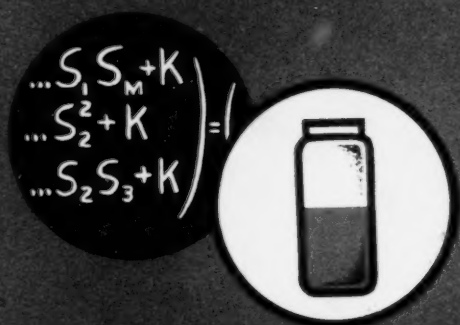
## ABSTRACT

This booklet surveys all the research studies (approx. 205), undertaken at ETS from July 1968 through June 1969. It offers research results to interested, nonresearch professional readers, and supplements more formal reports for professionals interested in research processes as well. Four major themes in the ETS research program are emphasized: (1) the assessing of human beings as individuals, rather than as "averages"; (2) assessing educational processes in terms of systems, rather than concentrating solely on separate processes; (3) a concentration upon research aimed at gaining an understanding of how children learn during their earliest years; and (4) initiating and carrying out research in a real world with all its complexities. The contents are grouped into four categories: Psychological Research; Educational Applications; Theories, Methods, and Systems Research; and Publications and Speeches of ETS Staff members. (CK)

ED052258

TM 000 752

# RESEARCH AT ETS: PROJECTS AND PUBLICATIONS



ED052258

# RESEARCH AT ETS: PROJECTS AND PUBLICATIONS

PERMISSION TO REPRODUCE THIS COPY-  
RIGHTED MATERIAL HAS BEEN GRANTED  
BY

EDUCATIONAL  
TESTING SERVICE

TO ERIC AND ORGANIZATIONS OPERATING  
UNDER AGREEMENTS WITH THE U.S. OFFICE  
OF EDUCATION. FURTHER REPRODUCTION  
OUTSIDE THE ERIC SYSTEM REQUIRES PER-  
MISSION OF THE COPYRIGHT OWNER.

U.S. DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
OFFICE OF EDUCATION  
THIS DOCUMENT HAS BEEN REPRO-  
DUCED EXACTLY AS RECEIVED FROM  
THE PERSON OR ORGANIZATION ORIG-  
INATING IT. POINTS OF VIEW OR OPIN-  
IONS STATED DO NOT NECESSARILY  
REPRESENT OFFICIAL OFFICE OF EDU-  
CATION POSITION OR POLICY.

DAVID LOYE  
*Editor*

COPYRIGHT © 1970 BY EDUCATIONAL TESTING SERVICE. ALL RIGHTS RESERVED.



# PREFACE

Educational research is a fascinating field because its aim is, quite directly, to discover how to enhance the quality of man's life.

Through educational research, for example, we may discover how to preserve the fresh curiosity and delight in learning of the young child. We may learn what threatens this gift, and how to remove the threat. We may also discover how we may more surely rekindle minds of all ages that have been turned off by life and circumstance, how to lessen hatred and hostility, and to increase love and respect—for others, and for ourselves. These are root concerns of current educational research, including that conducted by Educational Testing Service.

To translate goals into action, over the past two decades ETS has built a substantial base for research. During 1969-70 over 80 full-time research specialists (including 50 PH.D.'s and 28 M.A.'s), with an additional professional support staff of 55, were concerned with over 200 ongoing projects. These projects involved basic research in testing and measurement theory, individual personality and social behavior, human learning and cognition, and studies of how individual growth occurs under a wide range of influences. Research projects also attacked specific, sizable educational problems. Directions of continuing importance, for example, are research to aid disadvantaged peoples, to improve the conditions for early learning in homes and schools, to broaden and improve higher education, and to put the power of the computer at education's service.

Most of the research program is carried out within the ETS facilities near Princeton, New Jersey. The main Research Center has 80 offices for research psychologists. It also houses computer facilities, experimental laboratories, workshops for constructing experimental devices, and an experimental classroom for preschool youngsters whose growth in learning is observed by ETS child study specialists. Nearby is a specialized library with 15,000 volumes and 600 professional journals, which are mainly used by resident research workers but are available broadly to scholars in the field.

This setting for research is unusual in social science. In effect, it grew out of a decision made at the time of the founding of ETS as a nonprofit testing organization in 1948. The decision was to carry on research that might benefit education, and to fund the effort by investing in it one-half of ETS's revenue above costs. Thus, as school populations and the use of educational tests rose during the fifties and sixties, new funds became available to educational research.

These funds were used to bring together a core research group that made contributions to measurement theory and educational research

## PREFACE

from the late forties to the early 1960s. In 1965, with the aid of a grant from the National Science Foundation, ETS built the new Research Center, and, aided by another sizable long-term grant from the U. S. Public Health Service, it was able to expand its staff and its objectives in basic psychological studies.

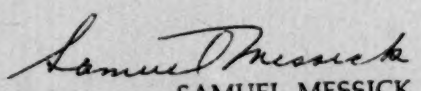
Thus, with ETS's own funds serving as the financial base, and grants from federal agencies, foundations, and other sources underwriting many specific projects, a productive research organization has been established over the years. The investment of time, money, and aspiration over two decades have also made of this Center an unusually stable setting for research—a crucial point, since good research requires sustained effort, and sustained effort (as anyone doing research knows) requires considerable stability of circumstance.

We trust the studies described in this publication will provide the reader with a convenient sampling of the effort that has been made to focus attention on the individual and his development.

*William W. Turnbull*  
PRESIDENT



May I direct your attention to the new publication *Research at ETS*. It describes recent and ongoing studies by ETS staff to those who look to research for answers to the problems of education and human development. The preface by William W. Turnbull gives a brief description of the background and resources for ETS research. My introduction provides an overview of the research effort and programmatic emphases. We would welcome your comments on this new publication.

  
SAMUEL MESSICK  
VICE PRESIDENT FOR RESEARCH  
EDUCATIONAL TESTING SERVICE

# CONTENTS

*Preface* 3

*Introduction* 15

## I. PSYCHOLOGICAL RESEARCH

### *Human Development*

- 24 EMMERICH The Parental Role: A Functional-Cognitive Analysis
- 24 EMMERICH A Developmental Study of Social Norms
- 24 EMMERICH Personality Development and the Concept of Structure
- 24 LEWIS Attention Distribution as a Function of Complexity and Incongruity in the 24-month-old Child
- 26 LEWIS Attentional Distribution as a Function of Novelty and Incongruity
- 26 LEWIS A Developmental Study of Learning Within the First Three Years of Life
- 26 LEWIS Developmental Study of Attention Within the First Two Years of Life
- 26 LEWIS Play Behavior in the Year-old Infant: Early Sex Differences
- 26 LEWIS Perceptual Cognitive Development in Infancy
- 28 LEWIS An Exploratory Study of Resting Cardiac Rate and Variability
- 28 LEWIS The Cardiac Response to a Perceptual Cognitive Task in the Young Child
- 28 LEWIS Cardiac Responsivity to Tactile Stimulation in Waking and Sleeping Infants
- 28 LEWIS The Meaning of an Orienting Response
- 28 WILSON The Magnitude of the Orienting Response in Children as a Function of Changes in Color and Contour
- 29 WOHLWILL Methodology and Research Strategy in the Study of Developmental Change
- 29 WOHLWILL A Longitudinal Study of the Formation of the Concepts of Conservation and Classification
- 29 BRACEY Memory Systems in Young Children

## CONTENTS

### *Learning and Cognition*

- 29 CARROLL Studies of the Acquisition of Grammar in a Foreign Language
- 30 CARROLL Factors in Comprehension of Verbal Material
- 30 CARROLL Learning from Verbal Discourse in Educational Media
- 30 CARROLL Application of the Lognormal Model of Word Frequency Distribution
- 30 JACOBS Training of Cognitive Functions in Children
- 31 FREEDLE Human Choice Behavior
- 31 FREEDLE General Choice Theory
- 31 FREEDLE Human Probability Learning
- 31 CAHEN Teacher Bias and Expectancies
- 31 CAHEN Estimation of Spelling Difficulties
- 35 CAHEN General Ability of Analytic Style
- 35 FREMER Recall of Approximations to English
- 35 HALE The Influence of Response Time and Stimulus Discriminability on Children's Dimension Preferences
- 35 HALE Interrelations Among Paired Associate Learning Tasks Involving Verbal and Pictorial Material
- 35 HALE Children's Retention of Paired Associate Learning Materials as a Function of Method of Presentation
- 36 KEENEY Children's Imitation of Utterances

### *Personality*

- 36 MESSICK Perceptual Factors in Cognition and Personality
- 36 MESSICK Personality Organization in Cognitive and Social Processes
- 36 MESSICK Desirability Judgments and Inventory Responses
- 37 MESSICK Judgmental Dimensions of Desirability and Frequency of Occurrence in the Prediction of Response Factors
- 37 MESSICK Judgmental Dimensions of Psychopathology
- 37 MESSICK Cognitive Styles and Psychopathology
- 40 MESSICK Cognition of Affect
- 40 MESSICK Dimensions of Perceived Item Similarity



- 40 FREDERIKSEN Learning to Formulate Original Responses in Relation to Anxiety and Defensiveness
- 40 WARD Creativity and Environmental Cues in Nursery School Children
- 40 WARD Rate and Uniqueness in Children's Creative Responding
- 41 DAMARIN The Psychology of Response Styles
- 41 DAMARIN Leveling-Sharpening as a Cognitive Style
- 41 DAMARIN Study of Objective Tests of Personality in Children
- 41 DAMARIN Study of Objective Tests of Personality in Adults
- 42 ALKER Personality Determinants of Multiple-Choice Item Preference
- 42 ALKER Are Bayesian Optimal Decisions Artificially Intelligent?
- 42 MULRY Personality and Test Taking Behavior
- 42 BLOXOM Test Anxiety Study

#### *Social Behavior*

- 43 KOGAN Negotiation Behavior Under Conditions of Observation and Consultation
- 43 KOGAN Individual and Group Decision Making Under Conditions of Responsibility for Others
- 43 KOGAN Determinants of Risk Taking for Others
- 43 KOGAN Effects of Anticipated Delegate Status on Group Risk Taking
- 43 CARLSON Cross-Cultural Study of Group Risk Taking
- 45 KOGAN Risk Taking in the Context of Intergroup Negotiations
- 45 KOGAN Pessimism-Optimism and Risk Taking
- 45 KOGAN Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma
- 45 HERMANN Negotiation in Leader and Delegate Groups
- 46 GULLIKSEN Intercultural Attitude Study
- 46 ROSENHAN Prosocial Behavior
- 46 ROSENHAN Observation of Prosocial Behavior
- 46 MESSICK Cognitive Styles and Attitudinal Consistency
- 46 STRICKER Dimensions of Conformity
- 51 TEGER Exploratory Research on the Escalation of Conflict

CONTENTS

- 51 SHORE Conceptual Complexity and Normative Social Expectations  
51 ZDEP Behavioristic Theory of Attitude Development

II. EDUCATIONAL APPLICATIONS

*Early and Primary Education*

- 54 CHITTENDEN Assessment and Development of Cognitive Skills in  
Young Children  
54 BALL Preschool Television for Children  
54 ANDERSON School Readiness Criteria Study  
54 KLING SEARCH  
58 IRONSIDE Literature Appreciation Test  
58 IRONSIDE Church School Curriculum  
58 MASSAD Reading Comprehension Ability, Writing Ability, and  
Linguistic Structure of the Written Language of Sixth-  
Grade Children

*Secondary Education*

- 58 HAVEN The College Board Candidate Survey  
59 PITCHER College Board Activity and Participation Statistics  
59 THOMPSON Curricular Appropriateness of the College Board  
Science Achievement Tests  
59 CARROLL U.S. National Testing in French as a Foreign Language  
59 DIEDERICH Tests of Cultural Development  
60 DIEDERICH Newspaper Tests  
60 BARROWS Foreign Language Innovative Curriculum Study  
60 TRISMEN Brevard County Project  
60 CASSERLY Students' Evaluation of the Advanced Placement Pro-  
gram and College Policy  
60 PETERSON The Exeter Alumni Survey  
61 HILTON A Study of Academic Prediction and Growth in Grades 5  
through 12  
61 HILTON Growth as a Function of the Amount of Training  
61 KATZ PSAT-AIM Norming and Follow-Up Study

- 61 PITCHER Follow-Up Study of a National Sample of High School Seniors, Phase Three

*Higher Education*

- 63 PETERSON The Scope of Organized Student Protest in 1967-1968
- 63 PETERSON Conversations Toward a Definition of Institutional Vitality
- 63 PETERSON Institutional Vitality Study
- 63 HARTNETT Trustee Study
- 64 DAVIS The Trustee in North Carolina
- 64 HARTNETT Agreement with Views of Martin Luther King, Jr., Before and After His Assassination
- 64 CENTRA Academic Growth in Predominantly Negro and Predominantly White Colleges
- 64 HARTNETT Differences in Selected Attitudes and College Orientations Between Black Students Attending Traditionally Black and Traditionally White Institutions
- 67 CENTRA Development of the Questionnaire on Student and College Characteristics
- 67 CENTRA College Freshman Attitudes Toward Cheating
- 67 HARTNETT Attitudes and Secondary School Backgrounds of Catholics Entering College

*Graduate and Professional Education*

- 68 BURNS Investigation of Graduate Admission Policies and Procedures
- 68 WARREN GRE Creativity Project
- 68 LUTZ Alternative Methods of GRE Advanced Tests
- 68 LANNHOLM GRE Cooperative Study of Predicting Graduate School Success
- 68 LANNHOLM The Use of GRE Scores and Other Factors in Graduate School Admissions
- 69 SCHRADER Validity of GRE Advanced Tests for Predicting Law School Success
- 69 LINN Continuation of Criterion Study: LSAT
- 69 SCHRADER Interpreting Performance of Foreign Law Students on LSAT and the Test of English as a Foreign Language

## CONTENTS

- 69 CARLSON Factor Analysis of the LSAT
- 69 FLESHER Summary of LSAT Research
- 69 COWELL Special Item Analysis of the Admission Test for Graduate Study in Business for Candidates Sponsored by the Consortium for Graduate Study in Business for Negroes
- 70 BOLDT ATGSB Grade Grouping Study
- 70 BOLDT Discrepant Predictors Study for the Admission Test for Graduate Study in Business
- 70 CROOKS Criterion Study: ATGSB
- 70 CARLSON ATGSB New Item Type Research
- 70 PITCHER ATGSB Moderator Variable Study
- 73 CROOKS In-Basket Study
- 73 SHARON The Moderating Effect of TOEFL on ATGSB
- 73 HARRELL Selection of MBA Candidates for Later Job Effectiveness

### *Cultural Diversity and the Disadvantaged*

- 73 ANDERSON Disadvantaged Children and Their First School Experiences: A Longitudinal Study
- 74 MEDLEY Classroom Variables
- 74 GORDON Planning for Innovation
- 74 CAMPBELL Test and Job Performance of Various Subgroups Within Specified Occupations
- 74 NORTON Evaluation of the Integration of Evanston District 65 Schools
- 74 MARCO Chicago Title I Evaluation
- 80 NORTON Objective Achievement Testing for the Evaluation of Chicago Title I Projects
- 80 ZDEP Evaluation of the Verona Plan for Sharing Educational Opportunity
- 80 BOLDT Miami Bilingual Student Study
- 80 ROSENFELD Vocational Development Study—Patterns of Negro-White Differences
- 80 FREEBERG Assessment of Disadvantaged Adolescents
- 81 SHARON Project Access Evaluation
- 81 STRICKER Dimensions of Social Stratification for Whites and Negroes

- 81 CARLSON Cross-Cultural Study of the Effect of Instructions on the Müller-Lyer Illusion
- 81 HILTON Home Environment and Educational Development
- 82 HILTON Socioeconomic Status and Academic Gain
- 82 EKSTROM Professional/Technical Assistance to the Head Start Research and Evaluation Demonstration Program

#### *Career and Vocation*

- 82 NOVICK Bayesian Guidance Technology
- 82 KATZ System of Interactive Guidance and Information
- 85 HILTON A Study of Intellectual Growth and Vocational Development
- 85 WARREN Vocational Development Follow-Up Study
- 85 FREDERIKSEN Effects of Organizational Climates
- 86 CROOKS The Executive Study
- 86 CROOKS In-Basket Development and Research
- 86 FLESHER Construct Validity of the CLEP Educational Psychology Subject Examination
- 86 IRONSIDE WGBH-TV Training Program for Law Enforcement Officials
- 86 GORDON Differential Behavior Patterns

#### *Evaluation of Educational Systems*

- 87 ANDERSON National Assessment of Educational Progress
- 87 TRISMEN Educational Systems Information Program
- 87 DYER Methods of Measuring School System Performance
- 87 SCHRADER Test Data as Social Indicators
- 89 HARMAN Validation Plans for Driver Education Training Programs
- 89 ROSENFELD Evaluation of a Summer Reading Institute
- 89 CAHEN Matrix Item Sampling

### III. THEORIES, METHODS, AND SYSTEMS RESEARCH

#### *Psychometric Theory*

- 92 LORD Strong True-Score Theory



# CONTENTS

- 92 LORD Project Mirkwood
- 92 LORD Tailored Testing Theory
- 92 NOVICK Statistical Theories of Mental Test Scores
- 92 NOVICK Statistical Models of Mental Tests
- 92 NOVICK Statistical Methods in Personnel Selection, Evaluation, and Guidance
- 93 NOVICK Bayesian Inference
- 93 JORESKOG A General Method for Analysis of Covariance Structures
- 93 KRISTOF A General Theory of Oblique Factor Rotation
- 93 KRISTOF On Proportional Profiles in Factor Analysis
- 93 KRISTOF Statistical Notes on Reliability Estimation
- 93 KRISTOF Structural Properties and Measurement Theory of Certain Sets Admitting a Concatenation Operation
- 94 KRISTOF A Theorem on the Trace of Certain Matrix Products and Some Applications
- 94 KRISTOF Estimation of True Score and Error Variance for Tests under Various Equivalence Assumptions
- 94 KRISTOF On the Parallelization of Trace Lines for a Certain Test Model
- 94 KRISTOF On the Sampling Theory of Reliability Estimation
- 94 SCHONEMANN On the Metric Multidimensional Unfolding
- 94 SCHONEMANN Scaling a Simplex Symmetrically
- 94 SCHONEMANN Fitting One Matrix to Another Under Choice of a Central Dilation and a Rigid Motion
- 95 SCHONEMANN On Two-Sided Orthogonal Procrustes Problems
- 95 HALL Variance Analysis System—MANOVA
- 95 ROCK Advanced Moderator Variable Study
- 95 BROWNE Gauss-Seidel Computing Procedures for Some Methods of Factor Analysis
- 95 BROWNE Precision of Prediction
- 95 BECHTEL Least Squares Matrix Factorization and Multidimensional Scaling
- 96 WINGERSKY Defining Groups in Multivariate Space

## *Methodological Research*

- 96 LORD Statistical Adjustments When Comparing Preexisting Groups

- 96 GULLIKSEN Learning
- 96 EVANS Independent Activities Questionnaire
- 97 EVANS Longitudinal Study of Early Predictors of Creative Achievements
- 97 HOLMAN Models for Measurement and Learning
- 97 BARROWS Error Choice Study
- 97 HILTON Cross-Sectional vs. Longitudinal Results
- 97 COFFMAN Review of Literature on Self-Fulfilling Prophecy
- 98 ROCK Decision Theory Approach to Multi-Stage Testing
- 98 SEIBEL An Investigation of Student Test-Taking Behavior
- 98 MARCO Cognitive and Affective Dimensions of Differential Achievement on Open and Closed Book Achievement Tests
- 98 MEDLEY Goal-Oriented Teaching Exercises
- 98 JORESKOG SCAT-STEP Factor Structure
- 98 DONLON Differential Weighting of Item Distractors in Data Sufficiency Items
- 101 OWEN Bayesian Methods of Item Allocation
- 101 SCHWARTZ Estimation of Test Weights by the Subjective Utility Function

#### *Computer Studies*

- 101 ANASTASIO Exploratory Project on Adaptive Programming
- 101 ANASTASIO Computer Generation of Test Items
- 103 LINN Computer-Assisted Testing
- 103 LEAF A Computer Model of Everyday Human Thinking
- 103 HARMAN Multivariate Statistical Analysis
- 103 HARMAN Computer-Assisted Testing Symposium
- 103 LAVINE Model for ETS Shipping
- 104 LAVINE Scheduling Techniques
- 104 BEATON Multivariate Change Pilot Study
- 104 LEAF A List Processing Extension to PL-I

#### IV. PUBLICATIONS AND SPEECHES OF STAFF MEMBERS

# INTRODUCTION

Over the past decade there has been rapid and exciting growth of educational research, and ETS research has been a significant part of that growth. This booklet attempts to describe its nature, plans, and potential benefits. It is a survey of all the research studies undertaken at ETS during one recent year, specifically all studies that were ongoing at any point from July of 1968 through June of 1969. Some 205 studies are reported here. They have been written and categorized primarily to convey information easily to a nonresearch professional readership interested in research results. However, they have also been written to supplement more formal reports in conveying information quickly to a professional readership interested in research processes as well as findings. At the end of each summary the reader will find either an ETS research bulletin number or a journal listing, whenever such complete information on a study is available.

Our intention in this introduction is to provide an orientation to this very large body of work by calling attention to some of its underlying purposes. In particular, four major themes in the ETS research program will be emphasized, and we shall attempt to relate a few exemplary studies to these themes in presenting them. The four themes are these: (1) the assessing of human beings as individuals, rather than as "averages"; (2) assessing educational processes in terms of *systems*, rather than concentrating solely on separate processes; (3) a concentration upon research aimed at gaining an understanding of how children learn during their earliest years; and (4) initiating and carrying out research in a real world with all its complexities.

## *Assessing Individuals Rather Than "Averages"*

Traditional questions in education and psychology have frequently spawned answers that are either downright wrong, in that they summarize findings "on the average" in situations where a hypothetical "average person" simply doesn't exist, or else are seriously lacking in generality, in that they fail to take account of the multiplicity of human differences and their interactions with environmental circumstances.

An example is the "horse race" question typical of much educational research of past decades: Is textbook A better than textbook B? Is teacher A better than teacher B? Or, more generally, is treatment A better than treatment B? Such questions are usually resolved by comparing average gains in achievement for students receiving treatment A with average gains for students receiving treatment B. But suppose treatment A is

better for certain kinds of students and treatment B better for other kinds of students? A completely different evaluation of the treatments might result if some other, more complicated questions had been asked, such as "Do these treatments interact with differences in personality and cognitive characteristics of students—or with differences in their educational history, or family background, or community, or culture—to produce differential effects upon achievement?"

This more complex type of question has been chiefly—though by no means exclusively—explored through studies that the reader will find grouped in this report under the categories of *personality*, *computer studies*, and to some extent *methodological research* and *cultural diversity and the disadvantaged*. In personality, work on cognitive styles has been particularly extensive at ETS. These are dimensions of individual differences in the performance of cognitive tasks that seem to reflect consistencies in the manner or form of cognition, as distinct from the content of cognition or the level of skills displayed in cognitive performance. A characteristic study reported in this booklet is that of Fred Damarin on leveling vs. sharpening as a cognitive style. Other studies reported here explore cognitive styles with such self-descriptive titles as scanning, breadth of categorizing, cognitive complexity vs. simplicity, reflectiveness vs. impulsivity, and tolerance for incongruity.

A promising adjunct to this type of interest is work in computer studies, such as Anastasio's exploration of adaptive programming, Linn and Rock's study of computer-assisted testing, and Lord's study of tailored testing theory. The long-range aim of such studies is to put the computer to use not merely as a great batch-processor of scores and other data, but also as an interactive aid to students, teachers, and guidance personnel as a means of selectively tailoring testing as well as instruction to the individual. Particularly noteworthy in this connection are Novick's study of Bayesian guidance technology and Katz's study of the feasibility of developing a system of interactive guidance and information for use in junior colleges.

### *Assessing Education in Terms of Systems Rather Than Disparate Parts*

In any consideration of the individual, one soon encounters the question of environmental influence. But something more than individual and environment is involved—the *interaction* of the two is critical. To deal empirically with this complex phenomenon of interaction, it has become mandatory to view human development, and hence education as well, in *systems* terms.

In this approach, personal characteristics and concerns are viewed in relation to social and environmental characteristics, and both are systematically combined with concerns about the effects of educational treatments. In this way a useful conceptual framework for educational



and psychological research may be developed that will help answer questions of how all the components interact and flow toward certain outcomes. Arising naturally from such a view, but more complicated than the "horse race" comparisons mentioned earlier, are such questions as "What dimensions of educational experience are associated with growth on dimensions of cognitive functioning, or with changes in attitude or affective involvement, and what social and environmental factors moderate these effects?" The need for such a multivariate interactional approach derives from the fact that we are dealing in education and psychology with a complicated system composed of many differentiated and overlapping subsystems, so that even in research on presumably circumscribed issues it is important to recognize the interrelatedness of personal, social, environmental, and educational factors.

Undoubtedly the largest example of educational system assessment with which ETS was involved during the time of this report was the National Assessment of Educational Progress. This is a project that seeks to establish for education an annual measure of national, system-wide progress comparable to the Gross National Product index for economic progress. Along with several other organizations, ETS is involved in various phases of the enterprise, including the specification of objectives, the construction of exercises, and the conduct of field trials.

Dyer's method of measuring school system performance and Trismen's educational systems information program are other notable examples, in the *evaluation of educational systems* section, of attempts to apply system-analytic techniques to education—in one case, the development of educational performance indicators for the New York State school system, and in the other the development of a computerized educational monitoring system generalized to fit the evaluation needs of a variety of school districts.

However, multiple influences upon behavior should not only be considered at the grand level of systems analysis, but at much simpler levels as well—such as in developing and evaluating a measure of academic achievement—where we sometimes forget that even specific responses are frequently complexly determined and buffeted by a host of environmental influences. Work with such measures as the Sequential Tests of Educational Progress (STEP) and School and College Ability Tests (SCAT) within the context of the ETS Growth Study (see *secondary education*) is an example of this kind of emphasis.

### *Research on Early Learning*

By the end of the time span for this report—that is, by June of 1969—ETS's involvement in research in early learning was fully a third of its total research effort, most of it supported by grants from foundations and government agencies. The need to take a systematic look at present preschool programs and to develop useful alternatives for this and later



age levels is widely recognized. However, not so well known is the related need for a systematic development of instruments for assessing the cognitive/intellectual and the personal/social dimensions involved in present or prospective programs. Without such measures, meaningful evaluative studies of experimental programs in early education are difficult, if not impossible, to carry through.

In the mid 1960s, ETS helped pioneer in this particular area of need with the new "Let's Look at Children" measures, based largely on the theory of Jean Piaget, that were developed for first graders in the New York City school system. Subsequently, the Carnegie Corporation has funded a project to adapt to preschool age needs this approach that "teaches the teacher" while carrying out combined assessment and instructional tasks with the children. A report of this research by Chittenden and others is found in the *early and primary education* section.

The greatest range of early childhood measures, however, is being used in the Longitudinal Study of Disadvantaged Children and Their First School Experiences (see *cultural diversity and the disadvantaged*). No less than 30 measures are included in the battery for this study, some of which are important modifications of time-tested measures whereas others are new measures developed in keeping with the findings of the most useful studies in this area over the past decade.

Also of considerable importance to the understanding of early learning is the broader area of developmental studies. It has become increasingly evident that we must do more basic research on the nature of the developing human being during his formative years in order to understand and uncover the major dimensions of human growth. Quite fittingly, studies of human development open this report—the work of Walter Emmerich, Michael Lewis, and (during a year's leave of absence from Clark University) Joachim Wohlwill on the methodology of developmental research. Emmerich's Developmental Study of Social Expectations is particularly interesting in this context for the way it illuminates how the behavioral expectations of those about us and our own expectations interact from our earliest years to help account for characteristic needs and behavior at each stage of life. Lewis' work makes use of highly sophisticated laboratory research technology to establish basic findings in the study of attentional processes and autonomic responses in relation to cognitive skills in infants.

### *Real Research in the Real World*

Is it possible to do real research in the real world? The answer is, "Of course"—but it's not easy. Not nearly as easy as doing real research in an artificial world, such as that provided by many laboratory settings. But our times require the frequent departure of the researcher from his laboratory into the field, and field studies require many compromises in "recommended" attitudes and procedures. A time of rapid social change

and swiftly emergent, often baffling, social needs requires an openness and a flexibility in researchers for which, unfortunately, few have been prepared by specific training. Universities and professors tend to perpetuate the techniques they know best and a vested interest grows in problems and approaches with which researchers feel comfortable. But neither the times nor its problems are comfortable, and so research is being forced to change to contend with the challenges posed by social change.

This more open and flexible kind of thinking is operative in many studies summarized in this report. In basic psychological studies its emphasis is perhaps best reflected by Kogan's risk-taking studies relating to problems of attaining world peace; studies by Ward and others in the nature and uses of creativity; Carroll's studies of our basic social tool, language; and Frederiksen's investigation of the effects of organizational climate on managerial requirements.

An important instance of the change in research emphasis from the laboratory to the field appears in the type of study done by the Educational Studies Division. The real-world needs of schools that prompted Title I program funding, for example, brought on a surge in the need for expert evaluation of educational experiments and innovations as well as of modified traditional programs. Most studies of this type are, in essence, very useful applications of traditional research techniques and experimental designs flexibly modified to meet the new problems of the times. Clustered here generally under *early and primary education*, *secondary education*, and *cultural diversity and the disadvantaged*, this research includes such studies as Ball's evaluation of the Children's Television Workshop series "Sesame Street"; Trismen's evaluation of an ungraded school system in Brevard County, Florida; Campbell's evaluation of the effects of bias in job selection and rating measures; Norton's evaluation of the integration of the Evanston, Illinois, schools; Zdep's evaluation of school busing in Verona, New Jersey; and Marco and Norton's evaluation of Title I programs in Chicago.

Another feature of this emphasis upon field studies in the real world is the attempt to develop new techniques and approaches relevant to real educational problems. Here one could mention Anderson's school readiness criterion study, Diederich's newspaper tests for social studies, Medley's pioneering work in developing new measures of teacher and classroom behavior, and Freeberg's development of new ability measures tailored to the interests and skills of teen-agers in the Neighborhood Youth Corps and similar programs.

A particularly successful phase of this type of effort is the work of Peterson, Davis, Hartnett, and Centra in developing measures designed to aid colleges in institutional self-study during a time of upheaval in higher education. Whether or not the reader is interested primarily in higher education, we would urge a perusal of the short reports in the higher education category describing the studies of institutional vitality, student protest, trustees, and the development of the questionnaire on

## INTRODUCTION

student and college characteristics.

Of all the ETS projects in the "real world," however, the most sizable—and quite possibly the most significant over the long range—is the Longitudinal Study of Disadvantaged Children and Their First School Experiences. Involving the thinking of a team of noted national experts in the area of early childhood, the efforts of a large task force of ETS researchers, and the active participation of mothers and other workers in black and white communities, this study brings together in one project much of what ETS research has learned from the past and aspires toward over the years ahead. The study, which began in 1968, plans to follow the educational progress of about 1,500 school children through 1974. We hope that summaries of Longitudinal Study findings will have much of use to report for American education.

*Samuel Messick*  
VICE PRESIDENT FOR RESEARCH

### *Note*

The source of funds that helped to support research projects summarized in the following pages is indicated, when appropriate, at the end of each summary. Publications resulting from the research are also indicated. Abbreviations are used in both cases.

These abbreviations represent the following funding sources: ARPA—Advanced Research Projects Agency, ATGSB—Admission Test for Graduate Study in Business, CEEB—College Entrance Examination Board, GRE—Graduate Record Examinations, HEW—Department of Health, Education and Welfare, LSAT—Law School Admission Test, NICHD—National Institute of Child Health and Human Development, NIMH—National Institute of Mental Health, NSF—National Science Foundation, OEO—Office of Economic Opportunity, ONR—Office of Naval Research, USOE—United States Office of Education, USPHS—United States Public Health Service.

These abbreviations represent the following publications: PR—Project Report, RB—Research Bulletin, RM—Research Memorandum, SR—Statistical Report, TDR—Test Development Report.

Requests for more information about any project should be addressed to the Division of Psychological Studies, Educational Testing Service, Princeton, New Jersey. When a Research Bulletin (RB) is listed, more information will be available, but it is often highly technical. When no RB is listed, it is quite possible that no additional information is currently available. In some cases, projects summarized here are works in progress, with publication of results still to come.

I

---

**PSYCHOLOGICAL  
RESEARCH**



## Human Development

### *The Parental Role: A Functional-Cognitive Analysis.*

WALTER EMMERICH

A favorite suburban type for novelists and sociologists is the middle class parent riddled with doubts about his or her capabilities as a parent. This study is based on the somewhat comforting thought that amid whatever familial chaos may exist, most parents hold reasonably definite beliefs about their own role in child rearing and about their degree of success or failure in the venture. To obtain parental beliefs of this type, a Parental Role Questionnaire (PRQ) was developed and given to 103 parents of children attending a university nursery school. Measures of parental attitudes and personality were also used. In this way, profiles were obtained of a parent's socialization goals (exactly what kind of person does the parent want his child to become), the parent's beliefs as to how these goals are best achieved, how well equipped he is for the socializing task, and how successful he feels he has been. One finding was that parents tended to favor direct methods to encourage desirable child behavior, and indirect methods to discourage undesirable behavior in the child. Perhaps most generally useful, however, were the patterns tying child-rearing beliefs with personality structure in the parent. This research should be of interest to behavioral scientists concerned with family processes, parent-child relationships, parental attitudes, role theory, cognitive processes in social behavior, and child development. NICHHD. RB-69-65. Also, *Monographs of the Society for Research in Child Development*, 1969, 34 (8, Serial No. 132).

### *A Developmental Study of Social Norms.*

WALTER EMMERICH, KARLA GOLDMAN, ROY SHORE

- We take for granted that most adults can distinguish between "appropriate" and "inappropriate" social acts. However, acquiring such understanding is a developmental process about which

little is known that extends throughout childhood and adolescence. This study examines social norms held by the self and those attributed to significant others (parents, peers, males, females). Studies with 680 child and adolescent subjects indicate that social norms vary, depending on whether: (1) the holder of the norm is the self or another, (2) the holder of the norm is an adult or child, (3) the norm is directed in action toward a male or a female, and (4) the norm is directed in action toward an adult or a peer. Perhaps most important is the finding that norms vary systematically according to the child's age, and in some instances also according to the child's sex. The research should be of interest to behavioral scientists concerned with children's social cognitions and their development. It charts "natural" developmental trends on which education related to social development might build. The study also relates the maturity of a child's social norms to various indices of cognitive growth—especially academic aptitude test performance. Findings are relevant to questions of the "generation gap," as the study details age changes in the child's choice of peers versus parents as models. NICHHD.

### *Personality Development and the Concept of Structure.*

WALTER EMMERICH

As with many other things, how you approach the study of human personality has much to do with what you find. This study examines three approaches psychologists use in terms of how they are similar, or different, in their view of the development of personality structure. The first approach, the *classical* view, sees personality as an unfolding or growth of structures according to a relatively fixed sequence of possibilities—Freud's oral, anal, genital stages, or the cognitive growth stage views of Piaget. A second approach, the *differential* view, sees personality development in terms of collections of attributes or traits fairly common to

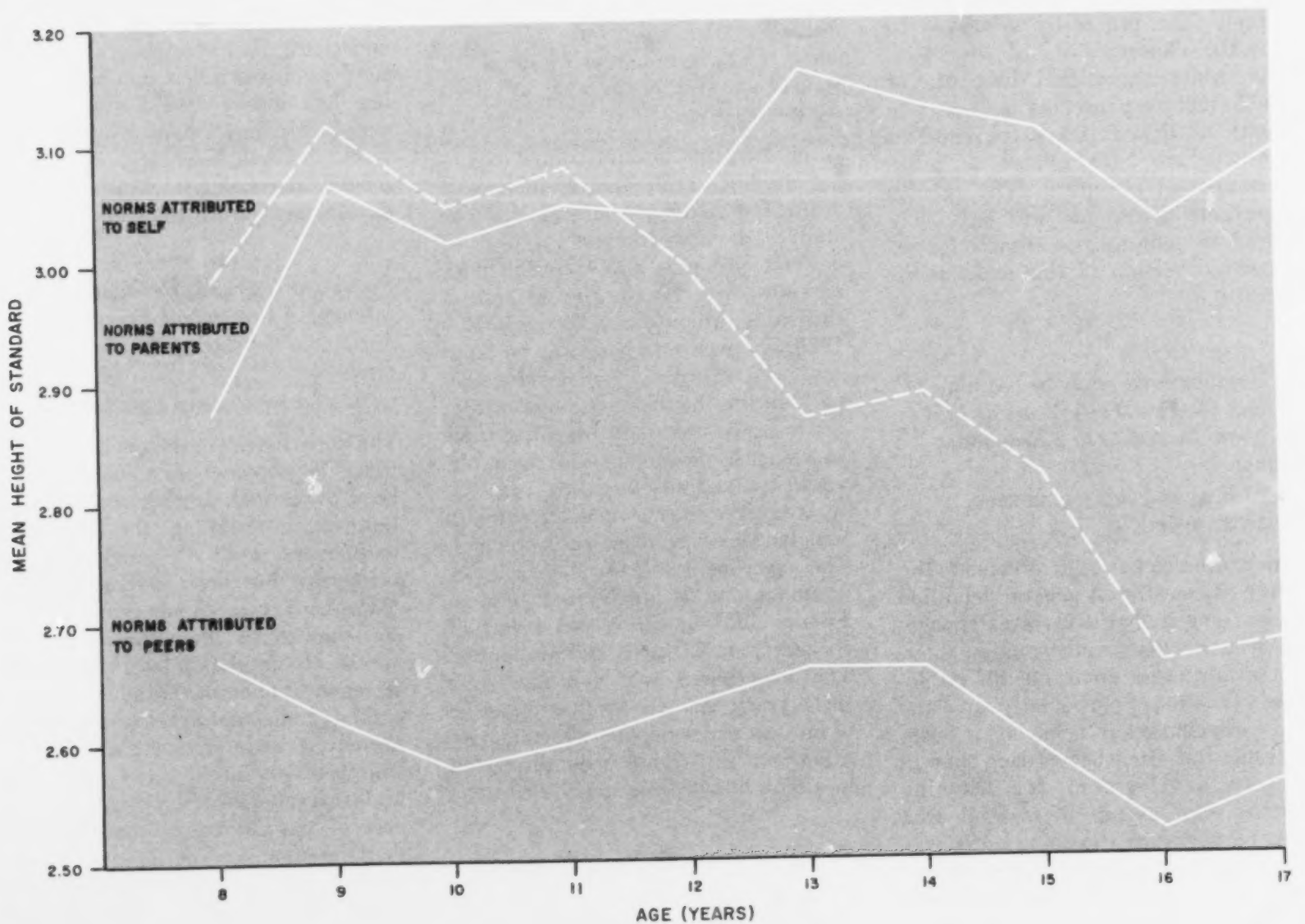
all of us, but wherein we differ greatly from one another—for example, we differ in intelligence, in need for achievement, and so on. Most testing is based on this approach, which was initiated by Galton in Darwin's time. A third approach, actually older in some ways than either of these but currently generating interest in new, potentially rigorously scientific forms, is identified in this study as the *ipsative* view. This approach is associated with psychologists such as Gordon Allport and examines personality structures in terms of individual rather than general development. NICHHD. *Child Development*, Vol. 39, No. 3, September, 1968, pp. 671-690.

### *Attention Distribution as a Function of Complexity and Incongruity in the 24-Month-Old Child.*

MICHAEL LEWIS, CORNELIA WILSON

Seeking to discover exactly why very small children pay attention to one thing rather than another, psychologists have tended to standardize a certain test procedure. A picture of a "jumbled" human face is used in comparison with a "normal" face, and complexity is tested with a more or less busy checkerboard pattern. This study varied that procedure to examine the effects on infant attention of using a jumbling of the whole human form as an incongruity test, and varying the number of figures presented as a test of complexity. The four pictures presented were, for example, (1) of three men, for a familiar, complex stimulus, (2) of a three-headed man, for an incongruous, complex stimulus, (3) of a single man, for a familiar, simple stimulus, and (4) of a man with his head upside down, for an incongruous, simple stimulus. The results indicated that both complexity and incongruity in the human form can affect attention distribution in the 24-month-old child. The research should be of interest to developmental, child, and experimental psychologists studying attentional processes. NICHHD.

As we grow from childhood into youth, three important determinants of our behavior are the norms we have for ourselves, the norms we believe our parents have, and the norms of our peers as we perceive them. This figure shows how the relationships of these three norm levels change as the child grows older, from age 8 through age 17. (These norms are for behaviors directed toward parents; see Emmerich, *A Developmental Study of Social Norms*.)



*Attentional Distribution as a Function of Novelty and Incongruity.*

MICHAEL LEWIS, CORNELIA WILSON

This study was concerned with seeing whether children pay attention differently to stimuli that are defined as novel as opposed to those defined as incongruous. The underlying purpose was to answer a theoretical question. From one position it is possible to predict not only that a familiar stimulus (a picture of a cat or horse, for example) would prompt less attention than an incongruous stimulus (a cat's head on a horse's body), but also that a novel stimulus (an anteater) would compel more attention than either of the other two. However, alternative theories could be used to support the argument that novel and familiar stimuli would be fairly equal in gaining a child's attention, and both would have less appeal than the incongruous stimulus. One difficulty lies in defining the stimuli. The procedure used was to rate the children's verbal responses. The results suggest that the nature of the verbal response, as well as the length of time it takes the child to respond, are both related to other measures of attention. The results were ambiguous, and will be chiefly useful in refining the stimuli for an improved version of this experiment. NICHD.

*A Developmental Study of Learning Within the First Three Years of Life: Response Decrement to a Redundant Signal.*

MICHAEL LEWIS, SUSAN GOLDBERG, HELEN CAMPBELL

Much research has been directed to the study of learning. A general definition of learning is that it involves changes in behavior that result from an interaction with the environment. As infancy is a time of particularly rapid and dramatic changes in behavior, it offers a fertile but often bewildering time in human development for learning studies. One avenue of research that can be useful in threading one's way through this maze is the study of response decrement or the slackening of response to a repeated visual signal. This study investigated response decre-

ment in children from infancy through three years of age. It found that response decrement follows a lawful developmental pattern within the first three years, increasing with age. (This phase of the study also suggested that this decrement represents a true learning phenomenon rather than a fatigue model.) It also found that response decrement in the young child correlates with and is predictive of other learning tasks and of individual differences in perceptual-cognitive capacities. In other words, the research suggests that how children respond to this single measure of response decrement may be predictive of their abilities in a wide variety of intellectual skills. The research should be of interest to experimental psychologists concerned with learning and child developmental processes. NICHD.

*Developmental Study of Attention Within the First Two Years of Life.*

MICHAEL LEWIS, CORNELIA WILSON, MARCIA HARWITZ, PEGGY BAN, NANCY VAN WYK

After a certain amount of initial exploring of a research area, it often becomes desirable to attempt a large, long-range study that may integrate a host of disparate findings into a useful block of knowledge. That is the purpose of this study, which got underway in May 1969. Its specific purposes will be (1) to study response decrement developmentally during the first two years of life; (2) to study cross-modality similarities in response decrement—for example, visual and auditory modalities; and (3) to study the effects of complexity in both modalities on response decrement. The experiment will use both a cross-sectional and a longitudinal design. Twenty children will be seen at each of six ages: 3, 6, 9, 12, 18, and 24 months. The experiment will continue over three years, and the results should be of interest to developmental, child, and experimental psychologists interested in attentional processes. NICHD.

*Play Behavior in the Year-Old Infant: Early Sex Differences.*

MICHAEL LEWIS, SUSAN GOLDBERG

Until recently the emphasis in studies

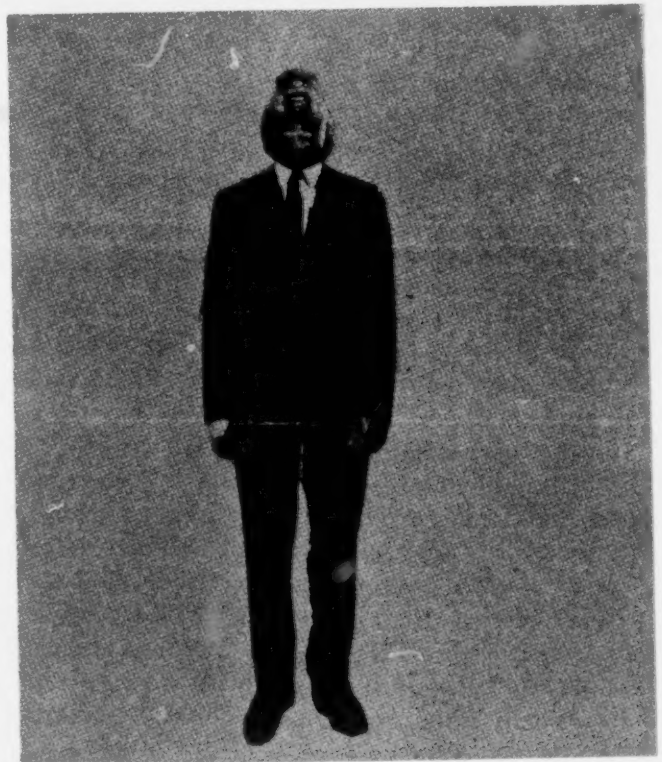
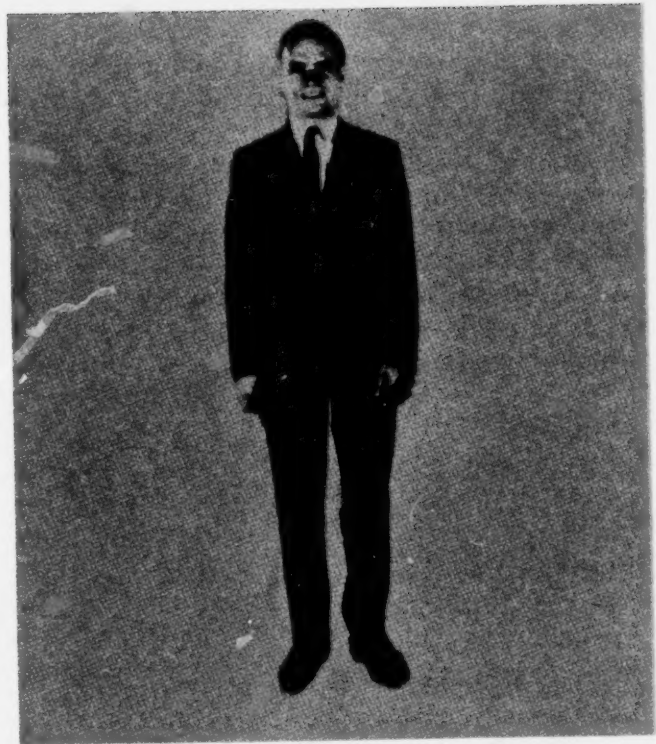
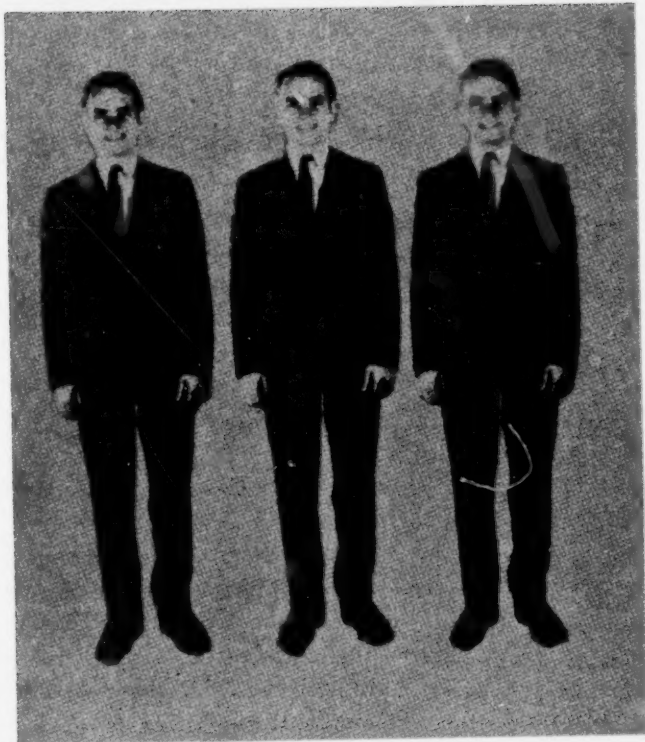
of child development was upon nursery and early grade school children. It has become increasingly apparent, however, that many earlier learning experiences are more important. The purpose of this study was to observe children at play, interacting with their mothers, and especially to note the effects of sex differences in this situation. Thirty-two boys and thirty-two girls, all 13 months old, were observed with their mothers in a standardized free-play situation. Striking sex differences in the children's behavior toward their mothers and in their play were noted. Earlier observation of the mothers' behavior toward the children when they were six months old indicates that some of these sex differences were related to the mother's behavior toward the children at this early age. The fact that such marked sex differences appear within the first year has important methodological implications for infant research. The findings emphasize the importance of checking for possible sex differences before pooling data and of considering sex as a variable in any infant research. The research should be of interest to experimental psychologists studying child development. NICHD.

*Perceptual Cognitive Development in Infancy: A Generalized Expectancy Model as a Function of the Mother-Infant Interaction.*

MICHAEL LEWIS, SUSAN GOLDBERG

The importance of mothers in the emotional development of children has long been recognized. During recent years, however, interest in the effects of mothers on a child's intellectual development has been swiftly growing. Out of this research has come the general finding that the amount and the variety of stimulation by the mother is an important parameter influencing the child's intellectual development. Additionally, it can be argued that a relationship between mother and child that leads to a generalized expectancy on the part of the child is more important than the amount of specific stimulation. Mothers and children were observed in a naturalistic laboratory situation. It was found that the following sequence seemed to be of paramount importance:





Incongruity and complexity of stimuli are two important variables affecting attention. These are four of the stimulus cards used by Michael Lewis to study how varying these components affects the attention of two-year-olds. (See Lewis, *Attention Distribution as a Function of Complexity and Incongruity in the 24-month-old Child.*)

if the mother responds *consistently* and *quickly* to the infant she seems to develop within him the expectation that he can influence his environment through his actions. This expectancy then seems to provide the motivation for additional interaction with his environment that, in turn, stimulates his cognitive development. The research should be of interest to child and experimental psychologists. NICHHD.

*An Exploratory Study of Resting Cardiac Rate and Variability from the Last Trimester of Prenatal Life through the First Year of Postnatal Life.*

MICHAEL LEWIS, CORNELIA WILSON,  
PEGGY BAN, MARCIA HARWITZ

There is evidence that differences in our autonomic nervous systems may be related to differences in cognitive abilities and personality. Specifically, the rate of our heart beat has been found to vary with our performance on a wide variety of tasks, as well as with personality differences. If this relation could be thoroughly documented through research, then it is possible that an infant's heart response might become a useful predictor of his probable development into childhood and even adulthood. This study was designed to see if any useful or normative relationships between heart rate and cognition and personality could be discerned from the period before birth through the first year of an infant's life. Measurements were made of the heart rate of 19 infants at regular intervals before and after birth. Despite the small sample, interesting developmental changes and promising consistencies were found, indicating the usefulness of further research in this area. The research should be of interest to developmental child psychologists and to physiologists. NICHHD.

*The Cardiac Response to a Perceptual Cognitive Task in the Young Child.*

MICHAEL LEWIS, CORNELIA WILSON

It has been found that a slowing down of our heart rate accompanies an intake of information and that our hearts beat faster during mental elaboration of this information, or when we wish to exclude

information. Using 44 one-month-old children, this study sought to find out whether the heart beat would slow down as the children solved a simple perceptual discrimination task, and also whether the degree of slowing down might be related to the accuracy of the response. There were several specific findings of some complexity. However, in general it was found that the degree to which the heart rate slowed down was significantly related to the child's general intellectual functioning as measured by a standard IQ test. Important and consistent sex differences in cardiac responsivity were also found—infant girls showed greater consistency of response systems, for example. The results suggest that, at least for very young girls, cardiac deceleration is directly related to both general intellectual capacity and efficiency. The research should be of interest to psychophysiologicals and child psychologists. NICHHD.

*Cardiac Responsivity to Tactile Stimulation in Waking and Sleeping Infants.*

MICHAEL LEWIS, CORNELIA WILSON,  
MARCIA HARWITZ

A harrowing problem for researchers is that of important but overlooked variables. Known as confounding variables, they can sometimes undermine the work of months and even years. A problem in the relatively new field of studying infant psychophysiological responses is that few studies have paid much attention to whether the infants were awake or asleep when tested. Hence, if there are clearly response differences between the two states, not only would this be a crucial variable to control in future work, but it could also render questionable much work in the past. This study sought to determine whether there are in fact cardiac response differences between waking and sleeping infants. Clearly discernible differences in response between the two states were found, indicating that sleeping or waking state differences must be accounted for in future studies. The research should be of interest to psychophysiologicals and developmental child psychologists. NICHHD.

*The Meaning of an Orienting Response: A Study in the Hierarchical Order of Attending.*

MICHAEL LEWIS, MARCIA HARWITZ

In the progressively finer differentiation of the important components of attentional processes, what is known as the orienting response (OR) has captured the interest of numerous researchers. The OR is involved in the tuning or focusing of attention that readies one for perceiving whatever it is that has caught one's attention. This study sought to reveal the inadequacy of present theory in this area, and to present alternatives. The specific questions explored were whether all discriminable stimulus changes produce the same degree of orienting response, and whether the degree of the OR depends on the salience of a particular stimulus. A visual signal was used, followed by variations of the signal that changed in color, size, number, and rotation. The results indicate a definite hierarchy of stimulus salience which is in turn, related to a subject's mental structure and conceptual development. That is, the different stimuli prompted different responses from different people, rather than the tendency toward a sameness that would have been predicted according to prevalent theory. The results are used to argue for a theoretical model of attending that takes cognitive principles into account in explaining the orienting response. The research should be of interest to experimental, child, and developmental psychologists interested in attentional processes. NICHHD.

*The Magnitude of the Orienting Response in Children as a Function of Changes in Color and Contour.*

CORNELIA WILSON, MICHAEL LEWIS

A difficulty with present theory for the orienting response is that it cannot be used to predict the magnitude of changes in the OR for various stimulus conditions. This study explored the effects of varying such stimulus dimensions as color and contour. A stimulus was shown six times, and then altered on the seventh presentation. The investigators were interested in how



accompanied by use of a "Language Master" (a device for presenting brief auditory samples along with printed or pictorial material). It was found that translations were superior to pictures in promoting grammar learning. It was also found that a confirmation rather than a prompting mode of practice was superior—that is, letting the child know he is definitely right or wrong as he makes each response, rather than prompting him to "discover" the right response after a period of some mystery. Besides their psycholinguistic relevances, the findings may have implications for the teaching of reading. The research should be of interest to foreign language teachers, especially those teaching young children, and to those concerned with the psychology of language learning. NICHHD. RB-68-43.

#### *Factors in Comprehension of Verbal Material.*

JOHN B. CARROLL, ROY O. FREEDLE

This program of research defines factors that may account for the way we comprehend spoken and written language. Materials include items from standard reading and listening comprehension tests. The presentation of stimuli is controlled by several advanced research methods, such as the use of compressed speech techniques. Three studies during the time period of this report made interesting progress. "Children's Imitations of Sentences Evaluated by Four Response Measures," by Roy Freedle, Terrence Keeney, and Nancy Smith, investigated pre-school age children's "telegraphic speech." Previous investigators had theorized that children omit function words and certain grammatical features because they have small memory spans and are not sensitive to these grammatical elements. This study suggests that these previous explanations are inadequate and probably wrong. Results are described in RB-69-79, and will be published in the *Journal of Verbal Learning and Verbal Behavior*. "Observations With Self-Embedded Sentences Using Written Aids," by Roy Freedle and Marlys Craun, explored the question of why people normally find self-embedded sentences with two or more

embeddings so difficult to understand. (A self-embedded sentence contains one or more sentences *embedded* within it; for example, "The woman whom the boy helped was tired" contains one embedded sentence.) Results are described in RB-69-22, and will be published by *Perception and Psychophysics*. "Some Interrelations Among Nouns: The Pursuit of Semantic Markers," by Roy Freedle, investigated questions of how categorization of words (nouns) allows one to discover the nature of semantic markers experimentally. Such markers have been proposed as one means by which we understand sentences. Freedle found two distinct groupings were produced by the experimental subjects, suggesting that if hierarchies of semantic markers exist, they presumably differ from individual to individual. The research should be of interest to psycholinguists and teachers of English and foreign languages. All three studies are supported by NICHHD.

#### *Learning from Verbal Discourse in Educational Media.*

JOHN B. CARROLL

As research in any area proliferates there generally arises a need to integrate findings and state what appear to be useful principles underlying many phenomena. Such a need has arisen in the study of verbal discourse, particularly because of the relevance of work in this area to the development of instructional methods employing modern technologies such as film, television, tape recordings, programmed instruction, and teaching machines. Uncovering principles that might guide developers of such devices could be extremely helpful at this time. This study reviewed some 1,500 references in the literature to identify major findings as well as major gaps in knowledge. An objective of the study, which is not complete yet, is the development of an appropriate theoretical orientation to the material. The completed study should be of interest to educators in general, including classroom teachers, textbook and curriculum writers, film and videotape producers, and specialists in the psychology of learning or the psychology of language. USOE. RM-68-2. RM-69-11.

#### *Application of the Lognormal Model of Word Frequency Distribution.*

JOHN B. CARROLL

One useful way of studying language is by determining word frequencies, or how often certain words will appear in characteristic expressions that are under study. This approach makes possible the portrayal of findings in terms of word frequency distributions that can be shown graphically or put into mathematical formulae. The purpose of this study was to gain a better understanding of word frequency distributions in order to be able to predict from them the size and diversity of vocabularies. A number of models of word frequency distribution have been proposed, including the mathematical transformation known as the lognormal distribution. However, the rationale for a lognormal distribution had not previously been developed. A mathematical model was developed, along with computer programs to test the model on extensive data. The most notable finding was that the lognormal distribution does provide an adequate fit to the data and can be used to predict vocabulary size and diversity. The study is relevant to education in that, in a general way, its findings could help define the vocabulary load in curricula in English and foreign languages. The findings also have applications in the stylistics of written composition. The research should be of interest to psycholinguists, mathematical psychologists, and educators concerned with the teaching of vocabulary either in English or in foreign languages. NICHHD. RB-69-90.

#### *Training of Cognitive Functions in Children.*

PAUL I. JACOBS

This study sought to find out to what extent children's intelligence can be increased by training. A notable finding was that the cognitive skill of double-classification can be taught in such a way that it transfers to new materials and is retained over at least four months. Another aspect of interest was that a taxonomy of logical relations was needed to evaluate attempts to teach the *cognition of figural relations*

quickly the child became bored with the repeated stimulus (or response decrement), and how readily he responded to the change on the seventh trial (or response recovery). The findings revealed more boredom with a repeated achromatic stimulus than with a chromatic stimulus but no differences in response decrement were evident for a comparison of curved line versus straight line stimuli. There was also greater response recovery for the change from black and white to color than the reverse, and greater recovery for the change from a straight line to a curved line stimulus than the reverse. The research should be of interest to child development researchers studying the orienting response. NICHHD.

*Methodology and Research Strategy in the Study of Developmental Change.*

JOACHIM WOHLWILL

For the same reason that histories are written to explain nations, and biographies to explain individuals, psychological researchers are concerned with the developmental view of man, which examines how we change over time. The problem of measuring change for the psychologist is so complex, however, that relatively little completely satisfactory developmental research is carried out. This study systematically approaches the problem for the researcher examining problems of methodology, research design, data analysis, and scientific inference. In particular attention is given to: (a) the measurement of changes along behavioral dimensions over extended periods of

time, (b) problems of longitudinal versus cross-sectional data-gathering models, (c) describing and analyzing the developmental function, (d) analyzing interrelationships among variables over the developmental course, (e) applying experimental methods to the study of developmental functions, and (f) integrating the study of developmental changes with the study of individual differences. The research should be of interest to developmental and research psychologists generally. RM-69-24. RB-69-40. A book on the subject is being written by Dr. Wohlwill.

*A Longitudinal Study of the Formation of Concepts of Conservation and Classification, as Related to the Development of Spontaneous Cognitive Activities.*

JOACHIM WOHLWILL

The work of Jean Piaget has transformed developmental psychology by providing a central body of theory and an array of concepts that have opened many avenues for research. This study was an attempt to relate fundamental Piagetian concepts to spontaneous childhood activities over a period of time. Specifically, the purpose was to assess changes in children's responses to tasks of conservation, class-inclusion and class-intersection over an 18-month period, and to relate these to changes in their approach to situations that call for spontaneous activities involving ordering, measuring, and classifying. Tests of 94 children who were in Kindergarten and Grade 1 at the start of the study indicate a relationship be-

tween conservation and measuring activities—a minimal level of proficiency in measuring seems necessary for advances in conservation to occur. Similar though less clear-cut relationships have been found between class-intersection and conservation. The research should be of interest to child and developmental research psychologists.

*Memory Systems in Young Children.*

GERALD W. BRACEY

There is no appreciable body of research on memory in young children. This project was a pilot study to remedy this situation by using a new technique for assessing rate of forgetting with a set of culture fair materials, which can be standardized for any culture. Other instruments were used to study the organization of memory. Not enough data were collected to yield conclusive findings; however, the following indications are noteworthy: (1) the forgetting curve was very steep, (2) organization of short term memory appears to be very rigid, and (3) failure to recall appears to be a problem of retrieving a stored item, and not a problem of storing the item. The research should be of interest to developmental psychologists, experimental psychologists interested in memory, and possibly teachers. It also has some bearing on memory work with adults. Further research of this type may be relevant to education in suggesting optimal rates for the presentation of information by teachers or machines.

## Learning and Cognition

*Studies of the Acquisition of Grammar in a Foreign Language.*

JOHN B. CARROLL

An unresolved question for psycholinguists is whether an "audiolingual habit formation" theory or a "cognitive code learning" theory is most ap-

propriate for guiding foreign language learning. That is, can foreign languages best be taught by the drilling of great numbers of words and speech pattern equivalents or by teaching a relatively smaller number of grammatical rules for making sentences? This is one of a series of studies aimed at finding

answers to various phases of this general question. Specifically, this study sought to discover whether children learn a foreign language more easily by seeing picture sequences that illustrate sentence patterns, or by translations into their own language. An individualized instruction method was used,



domain of intelligence. This taxonomy has been developed. Previous attempts to train intelligence have generally not clearly defined the training procedure nor specified the relations between operations used for training and those used for measuring intelligence. This approach takes care of both needs. The research should be of interest to developmental psychologists, curriculum builders, and geneticists. It is particularly relevant to the controversial conclusions of Dr. Arthur Jensen which, as he admits, apply only to presently known environments. As new manipulations of the learning environment are successful in increasing intelligence—as this study appears to have demonstrated—it may be expected that Jensen's pessimistic conclusions will be modified. The research also seems relevant to education in that, if intelligence can be trained, logically there should be a place for such training in the curriculum. The research is also relevant to the Head Start Program and generally to the issues of compensatory education. NICHHD. RB-68-65. RB-69-20. RB-69-88.

#### *Human Choice Behavior.*

ROY O. FREEDLE

A common complaint of educators is that for all the research that has gone into human learning we still know very little about it. Actually a very detailed knowledge of human learning processes has been accumulated over the years, but a great problem today is how to reduce this accumulation to something relatively clear theoretically—both to expose areas that need more research, and to communicate to educators what is actually known of learning. The goal of this project is to develop data-reducing theoretical capsules of this nature in a variety of experimental settings. Specific learning phenomena being explored include paired-associate learning, human probability learning, syntactic and phonemic recognition studies, and more general psychophysical sensory discrimination studies. One approach being used is the analysis of synthetic asymmetric confusion matrices using a modified input to the Shepard-Kruskal computer scaling pro-

gram. Research to date indicates that this program will recover the specially constructed "double" configurations which seem to account in part for asymmetric matrices. RB-70-40. NICHHD.

#### *General Choice Theory.*

ROY O. FREEDLE

A second phase of the general probe launched in "Human Choice Behavior" involves Luce's choice axiom. When this axiom is recast in terms of the proportion of time spent in making, say, manipulative or observing responses, it appears to provide a way of studying its applicability to an infant's attentional behavior. It also seems to provide a means of indirectly determining whether the infant perceives subgroupings of specially constructed multidimensional stimuli. A second aspect of general choice theory concerns a modification of Luce's theory in a different direction. This revision deals with a quantitative account of the effects of "irrelevant" stimuli and anchor stimuli on response probabilities for choices made from the available set of responses. When completed the research should be of interest to researchers in cognition and human learning. RB-69-89. NICHHD.

#### *Human Probability Learning.*

ROY O. FREEDLE

A third phase of the "Human Choice Behavior" study involves probability learning. Two studies tested an assumption based on the learning model known as the k-span learning model. The assumption in this case concerns the subject's memory of the last k events and the effects this information exerts on his subsequent choices. The evaluation of the memory assumption for k events indicates that a subject performs best on patterns of k events which involve few runs and performs worst on patterns involving many runs (with the exception of those patterns which involve a regular alternation of events). This indicates that the k-span learning model must be modified to account for such findings. RB-69-4. NICHHD.

#### *Teacher Bias and Expectancies.*

LEONARD S. CAHEN

For years it has seemed apparent to many observers that a teacher's high expectations influenced a student to greater performance while low teacher expectancies worked against a good showing by the student. Two fairly recent developments have spurred researchers to go beyond unsupported personal observations to try to establish—or disestablish—the phenomena. One goad has been the frequent contention of observers in ghetto schools that low teacher expectancies, or a negative bias, was depressing the performance of the black students in these schools. The other goad has been an increasing number of experiments which, on one hand, are establishing the phenomenon as indeed very much operative, but on the other hand, are raising a great many questions about exactly how it does and does not operate under varying circumstances. The purpose of this study is to construct a theoretical model for experimental testing. An extension is projected beyond studying elementary school teacher-pupil interactions to a study of secondary school teachers. One finding to date is that the amount of bias attributed to manipulated expectancies seems to be a function of the amount of information given to the teacher to establish the expectancy. When completed, the research should be of interest to educational and social psychologists and researchers.

#### *Estimation of Spelling Difficulties.*

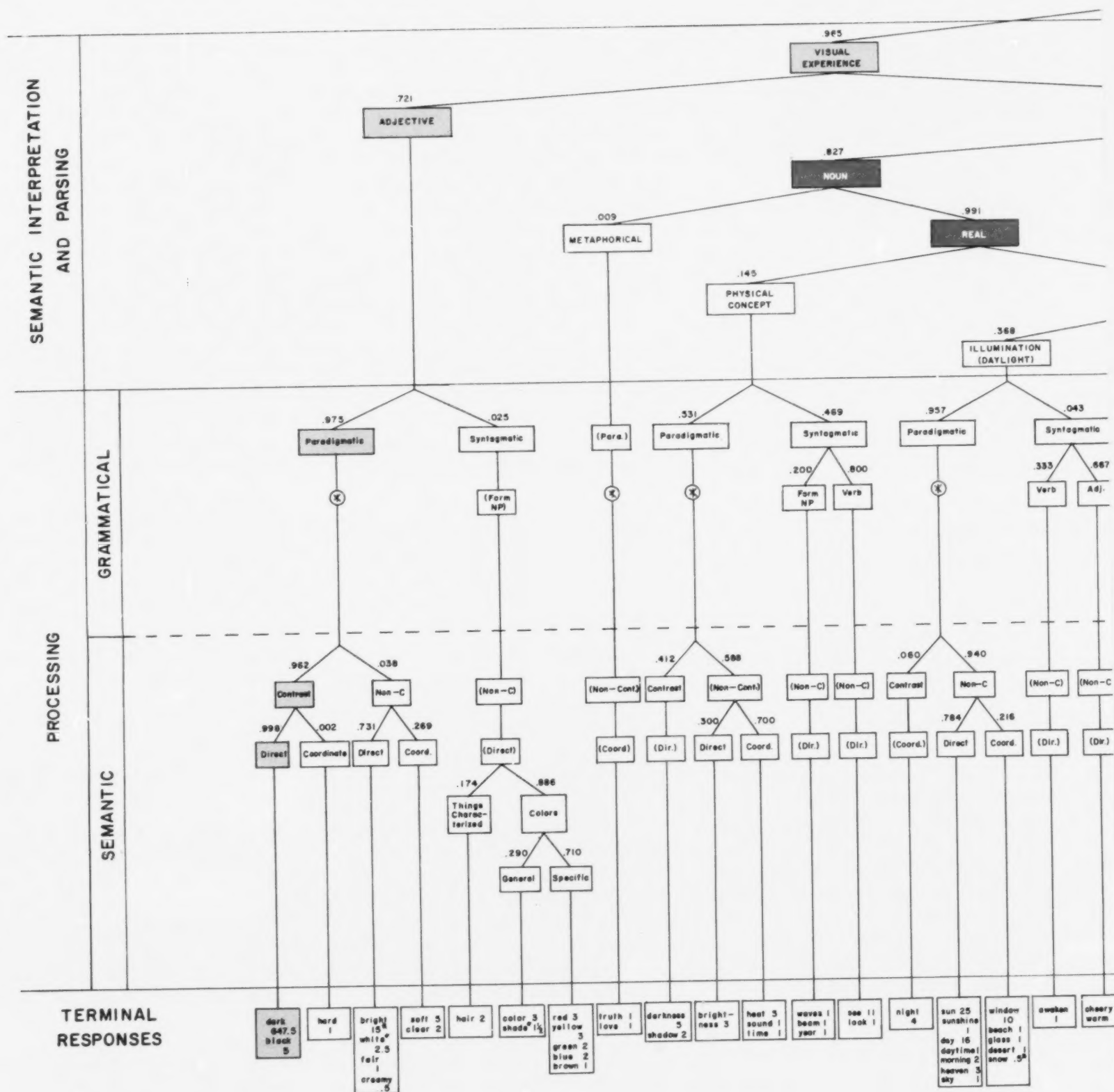
LEONARD S. CAHEN, MARLYS J. CRAUN,  
SUSAN K. JOHNSON

Spelling difficulties arise from numerous causes. This study is investigating spelling difficulty in terms of the stimulus properties of words and various cultural and psychological aspects of language. A review of the literature has led to the first phase of a model to predict spelling difficulty. The model involves a tentative list of 23 predictor variables, such as length of word, word complexity, and number of perceptual breaks in the word. The long-range purpose is to see how well the model will

(continued on page 35)

# TREE STRUCTURE OF ASSOCIATIONS TO LIGHT

How complex research in human learning has become is suggested by this chart. Shown are the hypothetical cognitive operations of two different individuals responding to the question "What word do you associate with the word

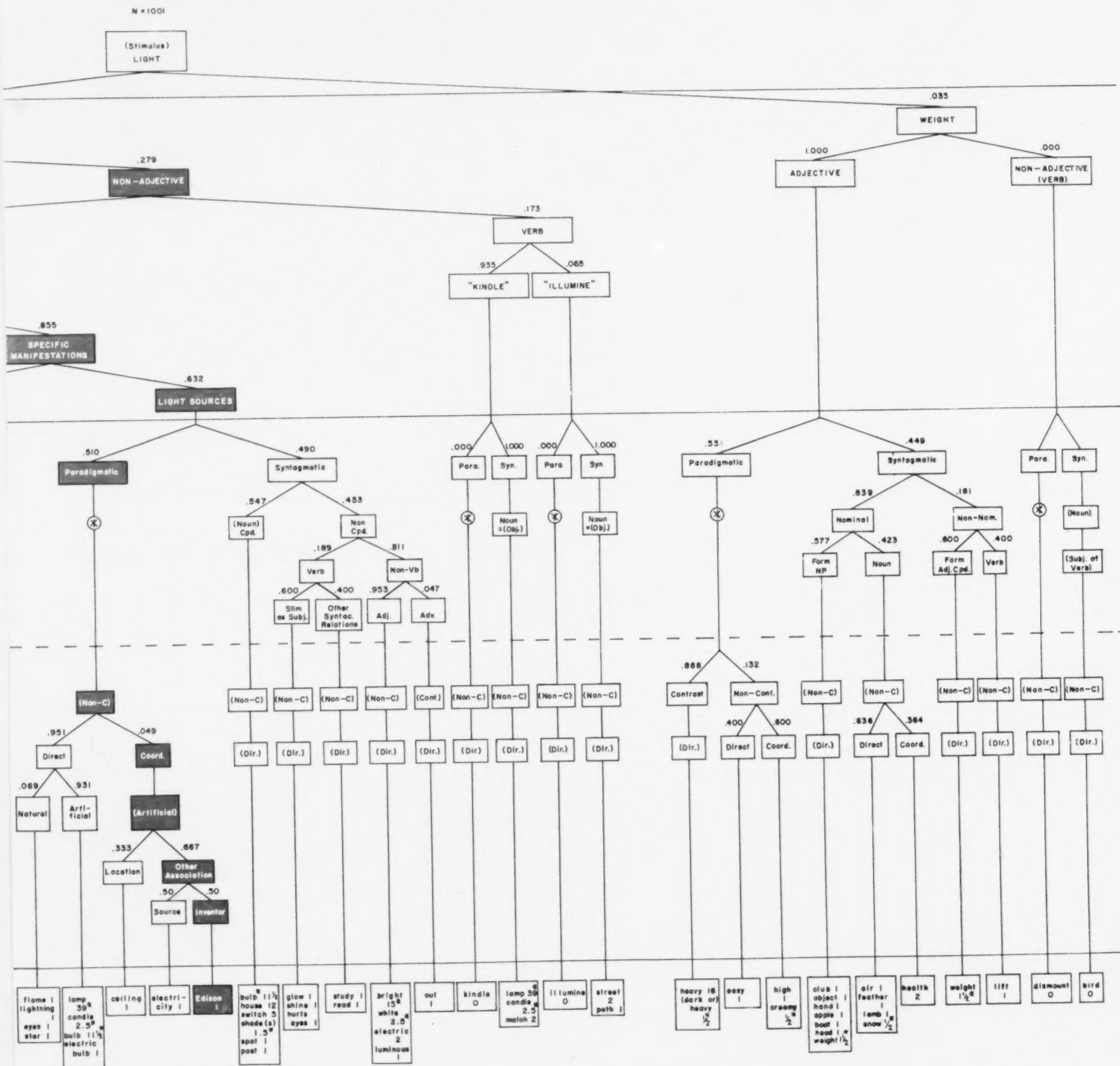


⊗ REDUNDANT - SAME PART OF SPEECH  
\* APPEARS IN 2 TERMINAL STRINGS - FREQUENCIES ADJUSTED



'light'? " Each person interprets this stimulus word (shown at the top of the diagram) as having a certain meaning and being a certain part of speech. Then the two may carry out quite different evaluations in terms of the grammatical and

semantic alternatives open to them. The hypothetical end result is a word-association response of "dark" in one case, and "Edison" in the other. (See Carroll, *Application of the Lognormal Model of Word Frequency Distribution*.)



```

C   PROGRAM GNLOGN GENERATES LOGNORMAL AT DEPTH D
      DIMENSION F(46), X(200), SPK(46)
      DOUBLE PRECISION DXI,CP
      REAL  NU, MUTAU, NT, MU, NTYPES
      INTEGER EN, F, CFK, IODD(2)/ZD549672A,ZE5C94642/
      CALL START ('*')
      CALL SETCLK
      Y1=-(CLOCK(A))/1000000.
      WRITE(6,999) Y1
999  FORMAT(9X,14HSTARTING TIME=,F14.6,8H SECONDS)
      10 YI=-(CLOCK(B))/1000000.
C   INST1=1, CONSTANT DEPTH; =2, POISSON DEPTH
C   INST2=1, TYPE DATA; =2, TOKEN DATA
      READ(5,25) LABEL1, LABEL2, LABEL3, EN, NU, D, MU, SIGMA, INST1, I
      INST2, INST3
      IF(EN.EQ.0) GO TO 1
      MUTAU=MU-SIGMA**2
      NTYPES=EXP(.5*SIGMA**2-MU)
      11 WRITE (6,19)
      12 WRITE(6,20) LABEL1,LABEL2,LABEL3,EN,NU,D,MUTAU,SIGMA,MU,NTYPES
      13 IF(INST1.EQ.1) WRITE(6,1013)
      14 IF(INST1.EQ.2) WRITE(6,1014)
      17 IF (INST3.EQ.2) WRITE (6,1017)
1017  FORMAT (25X,18HWITH CONSTANT P=.5)
      15 IF(INST2.EQ.1) WRITE(6,1015)
      16 IF(INST2.EQ.2) WRITE(6,1016)
1013  FORMAT(1H0,20X,27HGENERATED AT CONSTANT DEPTH)
1014  FORMAT(1H0,20X,35HDEPTH CONTROLLED BY POISSON PROCESS)
1015  FORMAT(25X,30HDATA ARE FOR TYPE DISTRIBUTION/)
1016  FORMAT(25X,31HDATA ARE FOR TOKEN DISTRIBUTION/)
      19 FORMAT(1H1,/,28X,2HEN,6X,2HNU,9X,1HD,7X,5HMUTAU,5X,5HSIGMA,8X,2HM
      1U,8X,6HNTYPES)
      20 FORMAT(4X,3A4,6X,I10,5F10.4,F12.2)
      25 FORMAT(3A4,I10,4F10.4,2X,3I1)
      IF (INST3.EQ.2) GO TO 28
      FNU=.5*((NU-1.)/NU)
      FNU2=1./NU
      28 SLNP=0.
      SLNP2=0.
      N=0
      IF(INST2.EQ.2) MUTAU=MU
      CFK=0.
      SP=0.
      SSPK=0.
      ID=1.00001*D
      IDP1=ID+1
      CEXPD=EXP(-D)
      DO 30 K=1, 46
      SPK(K)=0.
      30 F(K)=0
      Y2=-(CLOCK(C1))/1000000.
      ET=Y2-YI
      WRITE(6,1000) ET

```

---

Basic psychological research now frequently requires elaborate computer programming. This is a sample portion of the program (GNLOGN) for the Monte-Carlo generation of terminal distributions in a word-frequency distribution study by John Carroll. (See Carroll, *Application of the Lognormal Model of Word Frequency Distribution*.)

---

predict actual spelling difficulties, as reported by Greene's *New Iowa Spelling Scale* of 5,507 words. When completed, the research should be of interest to educational researchers. RB-69-72.

#### *General Ability of Analytic Style.*

LEONARD S. CAHEN

It has been found that how we approach or analyze problems may show consistencies according to the individual that can be characterized as *analytic styles*. The purpose of this study is to see how such styles relate to other cognitive and personality phenomena. The approach being used is based on Jerome Kagan's Matching Familiar Figures (MFF) Test, which has been found useful in identifying four kinds of analytic style: impulsive, reflective, quick responding, and slow responding. Students in the study have been asked to provide matchings that were the same as and matchings that were different from the stem or basic figure. Measures of the number of errors and response latency—or how long it takes them to respond—are being compared. A preliminary analysis of the data indicates that the dimension of latency is more stable within and across measures than the assessment of errors. The research should be of interest to researchers in cognitive and learning studies.

#### *Recall of Approximations to English.*

JOHN J. FREMER

Most of us take speech for granted, but to the psychologist the process whereby human beings dip into memory and bring up the right word or words in speaking is of a wondrous complexity. This study tried to uncover the relationship between the ability to recall unrelated words, and the ability to make use of contextual constraints as aids to recall. The comparisons were made with deaf children, and with hearing children from both disadvantaged and middle-class backgrounds. The data are still being analyzed; however, a preliminary analysis indicates that older middle-class children show an increased ability to make use of contextual constraint throughout the range used in the study. There appeared to be

a ceiling, however, to the level of contextual constraint that could facilitate recall by deaf children. The research should be of interest to those interested in language, memory, verbal mediation research, educational psychologists, and teachers of the deaf and the culturally disadvantaged. One educational implication is that it may be possible to identify children whose ability to use language cues as an aid to recall is poorly developed relative to their ability to recall unrelated words.

#### *The Influence of Response Time and Stimulus Discriminability on Children's Dimension Preferences.*

GORDON A. HALE

In many ways the eye is like a tunnel leading from the outer world into the inner world of our minds. Thus, the nature of its "gatekeepers" is of enduring interest to researchers. It is known that in carrying out their screening function, the perceptual "gatekeepers" evaluate the entering stimuli in terms of such dimensions as size, shape, color, texture, and so forth. This study sought to determine to what degree these evaluations are fixed and relatively inflexible, or whether they can be modified. A "dimensional preference" test was used whereby the presentation of stimuli was automated to carefully control both the duration of the stimuli and the time allowed the observer to respond to them. One finding was that to a high degree the tested kindergarten and third-grade children evaluated stimuli in terms of shape rather than color or other dimensions. This shape preference was also relatively fixed and hard to modify. The research is relevant to the problem of devising teaching materials that are based on an understanding of how children attend to certain features of the materials that are critical for learning. It should be of interest to educational and experimental psychologists.

#### *Interrelations Among Paired Associate Learning Tasks Involving Verbal and Pictorial Material.*

GORDON A. HALE

Research has found that learning pro-

ceeds more rapidly when the tasks involve pictorial rather than verbal materials, suggesting that the use of imagery facilitates learning. While previous research has indicated overall differences in the rate of learning for pictorial versus verbal materials, more needs to be known about the relationships between these two types of materials in relation to the learning processes. This study explored such questions and found a high correlation between two tasks involving pictorial materials—one abstract, and the other concrete. In contrast, it found a low correlation between two tasks involving verbal materials. The findings suggest a greater degree of diversity in the abilities required for learning verbal associations than in the abilities required for nonverbal learning. The research should be of interest to experimental psychologists and educational psychologists.

#### *Children's Retention of Paired Associate Learning Materials as a Function of Method of Presentation.*

GORDON A. HALE

This study was concerned with the effects of the method of presentation upon long-term retention in terms of interference theory. This is a long-standing theory which explains forgetting as a matter of old memories "interfering" with new memories, or vice versa. W. F. Battig, however, has proposed a hypothesis based on this theory that is somewhat counter to the theory's general thrust. It states that the amount of material children retain is influenced by the extent to which interference has been created among the items to be learned. Contrary to this prediction, this study found that conditions of low interference among items produced better retention than high interference conditions. The research should be of interest to educational and experimental child psychologists. It underlines the need to consider long-term retention as well as rate of learning as measures of learning, and is relevant to the development of programmed texts and computer-assisted instruction.



### *Children's Imitation of Utterances.*

TERRENCE J. KEENEY

Studies with young children comparing their imitation, comprehension, and production of speech generally show that their capacity for imitation is far superior to their ability to comprehend or produce speech. Under the assumption that the only sure measure of grammatical knowledge is comprehension, however, imitation has usually been dismissed as an uninteresting and probably irrelevant experimental task. This study sought to determine whether

a child's imitation of various forms of speech might, in fact, tap or make contact with his grammatical knowledge after all. Four-year-old children imitated strings of nonsense syllables (to see if lexical meaning was being tapped), strings of words (to see if grammatical meaning was being tapped), and grammatical sentences (as a control measure). The most notable finding was that imitation did indeed relate to the grammatical system of the young child, and to the lexical meaning system as well—the imitation of ungrammatical utterances was less accurate and slower

than the imitation of grammatical utterances. The significance of this finding is that it seems to indicate that imitation might be used as a measure of grammatical knowledge. This in turn opens the possibility of studying the development of the child's grammar at very early ages (1½ to 3 years) when comprehension measures are difficult to obtain and recordings of spontaneous speech underestimate a child's grammatical knowledge. The research should be of interest to psycholinguists, child psychologists, and educational psychologists. NICHD. RM-69-12.

## *Personality*

### *Perceptual Factors in Cognition and Personality.*

SAMUEL MESSICK, JOHN W. FRENCH

This study explored how our perception relates to thinking and feeling, or more technically and specifically, how individual differences in the speed and flexibility of closure in perception relate to cognition and personality. It is known, for example, that those who quickly report seeing definite shapes in perceptual tests and then hang fixedly to these first perceptions in the face of changing evidence often have somewhat different personalities and ways of thinking from those who are slower to gain closure and more flexible in their reports. In this study experimental tests of verbal and cognitive closure and of flexibility in reasoning and problem solving were developed. These measures were included in a battery of personality and ability measures given to more than 500 cadets at the U. S. Naval School of Aviation Medicine in Pensacola, Florida. A factor analysis isolated 14 factors relevant to the investigation in terms of speed or flexibility of closure in three performance areas: perceiving, thinking, and verbal expression. By means of statistical manipulation, these 14 factors were then reduced to four second-order factors: verbal comprehension, general reasoning, figural articulation (that is,

facility in structuring and differentiating figural stimuli), and symbolic articulation. The research should be of interest to personality researchers. ONR.

### *Personality Organization in Cognitive and Social Processes.*

SAMUEL MESSICK

This project is an attempt to deal with the interrelationships of personality, cognition, and social processes through a series of integrative studies that build from fairly specific projects to larger, more inclusive investigations. The studies range from fairly limited experimental investigations using only a few measures and relatively small samples of subjects to large integrative studies involving large samples of subjects, each of whom was administered more than 18 hours of testing. Something of the complexity and size of the project is perhaps best indicated by simply listing study areas, co-investigators, and publications. The studies were: "Categorizing Styles and Cognitive Structure," Nathan Kogan (*Perceptual and Motor Skills*, 1963, 16, 47-51, and 1965, 20, 493-497). "Cognitive Complexity: Dimensions of Role Constructs," Nathan Kogan (*Multivariate Behavioral Research*, 1966, 1, 165-175). "Leveling-Sharpening and Assimilation Tendencies," Fred Damarin and

Virgil McKenna (RB-68-28). "Tolerance for Unrealistic Experiences," Norman Cliff. "Scanning and Focusing," Harold Schiffman and Diran Dermen. "Cognitive Styles, Defenses, and Eye Movements," Lester Luborsky, Harold Schiffman, and Diran Dermen. "Style and Persuasibility," Fred Damarin. "Sex Differences in Cognitive and Personality Structure," Diran Dermen. "The Stability of Cognitive Styles and the Correlates of Change," Diran Dermen. "Dimensions of Analytic Attitude," Ferdinand Fritsky, Douglas Jackson, and Charles Myers, (*Journal of Personality*, 1963, 31, 346-370; *Educational and Psychological Measurement*, 1964, 24, 177-192; *Journal of Abnormal and Social Psychology*, 1964, 69, 313-318). "Cognitive Interference and Flexible Control." NIMH and NICHD.

### *Desirability Judgments and Inventory Responses.*

SAMUEL MESSICK

In the Thurstone method of constructing attitude scales, a number of experts are asked to judge the degree to which various statements reflect a particular attitude. An important finding some years ago was that the judges varied considerably and reliably in the ways in which they rated the various statements, the original study being of differences between judgments by



Negroes, pro-Negro whites, and anti-Negro whites. Since then a number of studies have explored the possibility that such judgments may prove superior to self-report measures for assessing personality—that they would reflect the personality of the rater as well as the properties of the items being rated. This study compared desirability judgments with personality inventory responses on the same items to ascertain the degree of possible overlap. That is, the subjects were asked at one time to judge the items for desirability and at another time to respond to a questionnaire that used the same items in the customary way. Correlations between desirability judgments and inventory responses were found to be sizable and systematically related to the level of the judged item desirability. Specifically, items judged to be extremely desirable or undesirable showed low correlations with inventory responses, whereas items within a broad neutral range of desirability showed substantial correlations with the responses. Factor analyses indicated that both similarities and differences exist between dimensions of judgment and dimensions of response. Judgment-response correlations were computed separately for each subject, with the finding of significant correlations with scales measuring defensiveness and the tendency to respond desirably. The research should be of interest to personality researchers and test developers. RM-64-13. USPHS.

*Judgmental Dimensions of Desirability and Frequency of Occurrence in the Prediction of Response Factors.*

SAMUEL MESSICK

The purpose of this study was to solve two interrelated problems: one, the identification of the components of the largest factor of personality inventories such as the Minnesota Multiphasic Personality Inventory (MMPI); and, two, evaluation of the use of consistencies in judging item properties as possible sources of valid data about characteristics of judges. The first problem was approached by considering judged desirability and two alternative connotative properties of items, the judged frequency of occurrence of the trait

represented by the item and the judged frequency of endorsement of the item. The second problem was approached by evaluating the degree to which individuals judged items differentially under alternative instructional sets. The results of this complex study should be of interest to personality researchers and are described in the article, "A Distinction Between Judgments of Frequency and of Desirability as Determinants of Response," in *Educational and Psychological Measurement*, 1969, 29, 273-293. The study was done in collaboration with D. N. Jackson of the University of Western Ontario. USPHS.

*Judgmental Dimensions of Psychopathology.*

SAMUEL MESSICK

When consistent individual differences in judging the desirability of personality statements were first uncovered, it seemed that these judgmental consistencies might serve as personality variables in their own right; i.e., that individual viewpoints about the desirability of certain personality characteristics might indirectly reflect consistencies in the personality or social values of the judges. This point has considerable methodological import; it implies that the seemingly objective task of rating the desirability of a trait might provide valid data for assessing the personality of the rater. In the present study an analysis of judgments of the desirability of MMPI items was undertaken to ascertain the relative number and nature of consistent individual viewpoints of the desirability of psychopathological characteristics. A sample of 150 college men and women rated and later responded to the 566 items. By examining these desirability judgments and the self-descriptive responses to the same items it was possible to determine the extent to which the two response processes reflect common personality variance. Of 13 judgmental dimensions which emerged, the largest corresponded to the mean global desirability rating, and 12 others reflected such judgmental points of view as: denial of lack of somatic control, impulsivity vs. religious preoccupation, femininity, socially deviant atti-

tudes, worry, and timid cautiousness vs. masculine adventuresomeness. Some of the judgmental dimensions were highly related to self-reports to identical items while other dimensions were unrelated to self-report scales defined by these judgmental dimensions. Discussion highlighted the complexity of interpreting judgmental viewpoints as personality trait measures, and the role of defensiveness and other response styles as mediators in the two processes. The study was done in collaboration with D. N. Jackson of the University of Western Ontario. USPHS.

*Cognitive Styles and Psychopathology.*

SAMUEL MESSICK

Over two decades considerable research has been directed to identifying certain individual differences in the ways we organize our perceptions, memory, and thinking. These differences have come to be known as *cognitive styles* for roughly the same reason that we speak of a particular *style* of writing, or of artistic expression as being characteristic of one person and not of another. Recent studies of the perception and cognition of schizophrenics have indicated these differences may be particularly striking between groups of mental patients—for example, patients with different symptom patterns (paranoid versus nonparanoid), different levels of adequacy in life adjustment prior to illness (good versus poor adjustment), and different lengths of institutionalization (the acute short-term patient versus the chronic long-term patient). Because of the extent of these differences, a study of mental patient responses may clarify many questions about cognitive styles in general. Conversely, examining those groups using measures different from the customary clinical measures may cast new light on the nature of thought disorders. To these ends an extensive battery of perceptual, cognitive, and personality measures was given to a sample of 95 hospitalized schizophrenics. The battery included measures of processes of interest to cognitive style investigators such as field-independence, element articulation, form articulation, scanning, verbal closure, perceptual closure,

(continued on page 40)

|     |       | -0.2585 | -0.1869 | -0.1153 | -0.0436 | 0.0     |
|-----|-------|---------|---------|---------|---------|---------|
| 10" |       | -0.2944 | -0.2227 | -0.1511 | -0.0795 | -0.0078 |
|     |       | *****   |         |         |         |         |
|     | 0.69  | .       |         |         |         | 2       |
|     | 0.67  | *       |         |         |         |         |
|     | 0.65  | *       |         |         |         |         |
|     | 0.63  | ..      |         |         |         |         |
|     | 0.61  | *       |         |         |         |         |
|     | 0.59  | *       |         |         |         |         |
|     | 0.57  | .       |         |         |         |         |
|     | 0.54  | *       |         |         |         |         |
|     | 0.52  | *       |         |         |         |         |
|     | 0.50  | ..      |         |         |         |         |
|     | 0.48  | *       |         |         |         |         |
|     | 0.46  | *       |         |         |         |         |
|     | 0.44  | .       |         |         |         |         |
|     | 0.42  | *       |         |         |         |         |
|     | 0.40  | *       |         |         |         |         |
|     | 0.38  | ..      |         |         |         |         |
|     | 0.35  | *       |         |         |         |         |
|     | 0.33  | *       |         |         |         |         |
|     | 0.31  | .       |         |         |         |         |
|     | 0.29  | *       |         |         |         |         |
|     | 0.27  | *       |         |         |         |         |
|     | 0.25  | ..      |         |         |         |         |
|     | 0.23  | *       |         |         |         |         |
|     | 0.21  | *       | 2       |         |         |         |
|     | 0.19  | .       |         |         |         |         |
|     | 0.17  | *       |         | 4       |         |         |
|     | 0.14  | *       |         |         |         |         |
|     | 0.12  | ..      |         |         |         |         |
|     | 0.10  | *       |         |         |         |         |
|     | 0.08  | *       |         |         |         |         |
|     | 0.06  | .       |         |         |         |         |
|     | 0.04  | *       |         |         |         |         |
|     | 0.02  | *       |         |         |         |         |
|     | -0.00 | ..      | 8       |         |         |         |
|     | -0.02 | *       |         |         |         |         |
|     | -0.05 | *       |         |         |         |         |
|     | -0.07 | .       |         |         |         |         |
|     | -0.09 | *       |         |         |         |         |
|     | -0.11 | *       |         |         |         |         |
|     | -0.13 | ..      |         |         |         |         |
|     | -0.15 | *       |         |         |         |         |
|     | -0.17 | *       |         |         |         |         |
|     | -0.19 | .       |         |         |         |         |
|     | -0.21 | *       |         |         |         |         |
|     | -0.24 | *       |         |         |         |         |
|     | -0.26 | ..      |         |         |         |         |
|     | -0.28 | *       |         |         |         |         |
|     | -0.30 | *       |         |         |         |         |
|     | -0.32 | .       |         |         |         |         |
|     | -0.34 | *       |         |         |         |         |
|     | -0.36 | *       |         |         |         |         |
|     | -0.38 | ..      |         |         |         |         |
|     | -0.40 | *       |         |         |         |         |
|     | -0.42 | *       |         |         |         |         |
|     |       | *****   |         |         |         |         |
|     |       | -0.2585 | -0.1869 | -0.1153 | -0.0436 | 0.0     |
|     |       | -0.2944 | -0.2227 | -0.1511 | -0.0795 | -0.0078 |

0.0996 0.1713 0.2429 0.3145 0.3861  
 0.0638 0.1354 0.2071 0.2787 0.3503 0.4220  
 \*\*\*\*\*

5 Calm

This computer print-out reflects some of the complexities of modern personality research. The computer is programmed to process research data, to plot these data in relation to X and Y axes, and to print a diagram of the results. In this instance 600 high school students were asked to judge the similarity or difference of various emotions. The X and Y axes represent two dimensions of perceived similarity—pleasantness vs. unpleasantness of the emotion, and intensity of the emotion. The various emotions are then shown in relation to these two dimensions. The significance of the technique is that it will give the research psychologist a precise graphic picture of many relationships. (See Messick, *Cognition of Affect*.)

A  
neutral

Interested  
3

joyful  
1

Delighted  
7

G  
amazed

1  
Surprised

9  
Anxious

D  
Excited

0.69  
0.67  
0.65  
0.63  
0.61  
0.59  
0.57  
0.54  
0.52  
0.50  
0.48  
0.46  
0.44  
0.42  
0.40  
0.38  
0.35  
0.33  
0.31  
0.29  
0.27  
0.25  
0.23  
0.21  
0.19  
0.17  
0.14  
0.12  
0.10  
0.08  
0.06  
0.04  
0.02  
-0.00  
-0.02  
-0.05  
-0.07  
-0.09  
-0.11  
-0.13  
-0.15  
-0.17  
-0.19  
-0.21  
-0.24  
-0.26  
-0.28  
-0.30  
-0.32  
-0.34  
-0.36  
-0.38  
-0.40  
-0.42

\*\*\*\*\*  
 0.0996 0.1713 0.2429 0.3145 0.3861  
 0.0638 0.1354 0.2071 0.2787 0.3503 0.4220



extremity and confidence of judgment, category width, risk taking, leveling-sharpening, conceptual differentiation, ideational fluency, and verbal comprehension. Also included were items of case history information, a clinical personality inventory, and rating scales for symptom expression, properties of delusions, and suddenness of onset of illness. Analyses are underway to isolate consistent patterns of cognitive style-personality trait interrelations. The study is being done in collaboration with Drs. Julian Silverman of the Agnews State Hospital and Douglas Jackson of the University of Western Ontario. When completed, the research should be of interest to personality researchers, test developers, and clinical psychologists. NICHHD.

#### *Cognition of Affect.*

SAMUEL MESSICK

Though the complexities of our emotional life have been explored by legions of psychiatrists and psychologists with a therapeutic aim, comparatively little formal study has been directed to the relation of emotion to cognition—or to how what we *feel* influences what we *know* or *learn*. Among many unanswered questions are: How do we perceive and evaluate emotions in others and in ourselves? How are these views of emotion then related to the ways we approach and handle intellectual tasks? What relation do our views of our emotions have to our ideological commitments (or lack of same)? This study, done in collaboration with Dr. Silvan Tomkins, was an attempt to answer such questions. Measures of cognition and personality were given to over 600 high school students, along with a task requiring them to judge the similarities and differences among 18 names of so-called primary affects, such as fear, surprise, joy, and contempt. The resulting data are being analyzed statistically to see if characteristic ways of perceiving emotions can be identified and related to individual differences in cognition and other aspects of personality. NIMH and NICHHD.

#### *Dimensions of Perceived Item Similarity.*

SAMUEL MESSICK

The study investigated the number and

nature of dimensions of perceived similarity among personality statements in terms of judged similarities in likelihood of endorsement. The dimensions of judged endorsement frequencies were in turn related to perceived content and desirability consistencies. Twenty-three items from the Minnesota Multiphasic Personality Inventory were selected to span six dimensions found in earlier studies. All possible pairs of the 23 items were rated by 150 college students in terms of the similarity of each pair of items with respect to their probability of endorsement. Multidimensional scaling of the responses revealed four dimensions of viewpoint about item similarity. One important distinction appearing in these dimensions is the separation of desirable from undesirable item content. The study is being done in collaboration with Douglas Jackson of the University of Western Ontario. USPHS.

#### *Learning to Formulate Original Responses in Relation to Anxiety and Defensiveness.*

NORMAN FREDERIKSEN,  
STEPHEN P. KLEIN, FRANKLIN EVANS

How does anxiety relate to the production of ideas? One hypothesis is that anxiety inhibits the production of ideas—that the highly anxious person tends to censor his ideas and express only a few, whereas the less anxious person does not engage in such self-censorship and thus expresses more ideas. This study explored the issue by using a form of programmed instruction that provides feedback to the learner in the form of ideas others have generated in answer to the problems with which he is presented. The hypothesis was that if people are given a chance to see ideas of others that are considered acceptable, this will encourage the highly anxious ones not to censor their own efforts and to express more ideas. The procedure did indeed increase the quantity of ideas produced, although not their quality. However, contrary to the hypothesis, the effect was the same for people with both high and low degrees of anxiety. Anxiety was not directly related to the quantity of ideas; the least anxious people produced the greatest number, the highly anxious

people the next greatest number, and people with only medium anxiety the least number of ideas. The research should be of interest to researchers in creativity and will appear in the *Journal of Educational Psychology*. NICHHD.

#### *Creativity and Environmental Cues in Nursery School Children.*

WILLIAM C. WARD

During recent years there has been a surge of research activity on creativity within the educational context, but a great many open questions still remain. This study explored the question of what differentiates the creative from the uncreative child. Creativity tests that required the child to name as many ideas as he could that met a simple problem requirement were given to 53 nursery school children. Two of the tests were given in a barren room typical of experimental studies, and the children were divided into creative and uncreative groups. The third creativity test was then given to part of the group under similar cue-poor conditions, but to others in a cue-rich testing environment—that is, in a room that contained a variety of stimulating pictures and objects. The stimulating room had a pronounced effect in encouraging the production of ideas by children, but with this important difference: While the creative children increased their ideational production in the stimulating setting, the uncreative children did not. The conclusion from this research was that scanning the environment for task-relevant information is an important strategy characterizing the creative child in his search for solutions to problems. RB-68-58. *Developmental Psychology*, 1969, 1, 543-547. NICHHD.

#### *Rate and Uniqueness in Children's Creative Responding.*

WILLIAM C. WARD

Creativity is commonly viewed as a special ability that has been given in notable quantities to some but not to others. This everyday view has its analogue in research among those who see differences in creativity from person to person as lying primarily within a range of special cognitive abilities. A counter view stresses personality



and motivational differences—in other words, whether we are creative or not is seen as depending more on interest and drive than on ability, in the narrow sense of specific cognitive abilities. To put the matter to a test, three ideational fluency measures were given to 34 seven- and eight-year-old boys. In keeping with other studies, it was found that over time the children gave fewer but better responses, *but that creative and uncreative children differed neither in the rate of responses given nor in the proportion of unusual responses*. The one major difference between the two groups was that the creative group persisted in generating responses, a high proportion of which were unique, after the uncreative group had stopped responding. The findings tend to support the view that differences in creativity may be more a matter of personality and motivation than of basic differences in cognitive abilities. The research should be of interest to educators and research psychologists interested in creativity. RB-68-36. *Child Development*, 1969, 40, 869-878. NICHHD.

#### *The Psychology of Response Styles.*

FRED L. DAMARIN, SAMUEL MESSICK

When people respond to self-report instruments (personality questionnaires, attitude and interest inventories), it is usually assumed that they are responding to the *content* of the items and that their responses are a function of the attitudes, opinions, or personality traits the instruments are intended to measure. Research, however, has indicated that there are pervasive general response tendencies, or *response styles*, that affect scores obtained on such measures. Examples of such styles are the tendency to respond to items in terms of their perceived social desirability, the tendency to agree with relatively neutral statements regardless of their content, and the tendency to use extreme categories when responding to rating scales. While these response styles have been treated chiefly as nuisance variables that contaminate tests, the present project aims at learning whether they can be explained as personality traits. A survey was made of correlational and factor analytic studies containing response styles measures and objective

(or performance) measures of personality. The review suggested a number of ways in which response styles research may be improved—specifically, that current measures of the styles may not be optimal, and that there are too few conceptually appropriate and methodologically sound criterion measures of the personality traits supposedly responsible for stylistic responses. More recent work suggests that response styles may be viewed as facets of the larger problem of self-descriptive behavior. A model for predicting the emergence of different response styles in children at different levels of general intelligence has also been developed. These studies should be of interest to personality researchers and all concerned with the validity of self-report instruments, and to researchers in psychotherapy and forensic psychology. RB-65-10. RB-69-10. NICHHD.

#### *Leveling-Sharpening as a Cognitive Style.*

FRED L. DAMARIN, VIRGIL MCKENNA, SAMUEL MESSICK

Though leveling-sharpening has gained acceptance as a testable personality trait, the starting point for this study was the question of whether such a trait actually exists. In general, leveling-sharpening is assumed to be revealed in tendencies to minimize (level) or maximize (sharpen) differences. A number of tests assumed to measure leveling-sharpening were used, among them tasks designed to assess the way the individual assigns numbers to stimuli in psychophysical judgments. It was found that there are probably several factors (rather than one personality trait) that merit the name "leveling-sharpening," but that none of these factors or the trait itself is well measured at the present time. It was also found that individual differences in judgment styles do not correlate very highly across different sorts of judgments and stimuli. It was found that memory style—for example, a preference for leaving things out versus a preference for embellishing memories—does show some generality. The research should be of interest to researchers in personality and perception. It is related to the general body of work on cognitive styles. Memory

styles findings reported in this study may also be relevant to the interests of educational researchers. USPHS.

#### *Study of Objective Tests of Personality in Children.*

FRED L. DAMARIN

An important problem in personality research is determining the ways in which personality changes, as well as remains the same, over a lifetime. One difficulty in research of this type is the equating of personality tests, so that the tests given a child may measure comparable traits in an adult. This study explored the question of whether personality traits identified in an adult can also be identified with the same sorts of tests in nursery school children. One general finding was that some traits can be identified in both children and adults with roughly the same test, but it isn't easy. A problem is that at very young ages the scores from performance tests of personality are very likely to be affected by the subject's intellectual level. Technically, the study is of particular interest in its use of a hierarchical factorization approach and its study of marker intercovariances in groups of different mental ability and in the techniques used to interpret factors. The work should be of interest to personality researchers and developmental psychologists. One relevance to education is the way the study suggests that children who succeed in school and those who fail may have quite different alternatives for further personality development. The study, co-authored with Raymond B. Cattell, is reported in *Monographs of the Society for Research in Child Development*, 1968, 33, No. 6, (Serial No. 122). NICHHD.

#### *Study of Objective Tests of Personality in Adults.*

FRED L. DAMARIN

This study sought to determine whether people with character disorders serious enough to require special social control (imprisonment, for example) differ from college students on various factors diagnosed by Raymond Cattell's performance tests of personality. It found that prisoners and other students differ on factors that represent skills and habits that are taught in school.

However, this could be accounted for by the fact that those with character disorders were maladjusted in school and often dropped out. And other discernible differences may have arisen from the situational differences—that is, the effects of being imprisoned versus the situational effect of being a student. The study should be of interest to personality researchers, psychiatrists, penologists, and those interested in new personality test designs. The findings may be relevant to education in that those with character disorders were markedly deficient in many skills associated with creativity. This suggests that the capacity to be creative as an adult may depend on *exposure* to schooling, even though it may not depend on extraordinary school achievement. It is also relevant to penology in suggesting the need for more research on the pathological effects of imprisonment.

#### *Personality Determinants of Multiple-Choice Item Preference.*

HENRY ALKER, JULIA CARLSON,  
MARGARET HERMANN

This study investigated the claim made by some critics that multiple-choice tests allow the "superficial" person adept at test taking to prosper while penalizing the "deep" or original thinker. A sample of 115 male students from a large Eastern university participated. Personality tests were used to classify these students either as "superficial" or as "deep" thinkers. The students were then given special multiple-choice tests developed from the verbal section of the College Board Scholastic Aptitude Test and English Composition Test. Results indicate that superficial thinkers actually do worse than deep thinkers on multiple-choice tests—at least, on the type of test used. Being adept at test taking tends to help deep thinkers more, while being able to recognize ambiguity tends to favor superficial thinkers. The research should be of interest to testing critics and developers. NICHHD. *Journal of Educational Psychology*, 1969, 60, 231-243.

#### *Are Bayesian Optimal Decisions Artificially Intelligent?*

HENRY ALKER, MARGARET HERMANN

When Bayesian formulas are used to predict human behavior, quite frequently the actual behavior will turn out to be more conservative than the prediction. This study explored two alternative hypotheses as to why this happens. The first hypothesis is that the Bayesian model is a good simulation of the information processing of intelligent, abstractive, rational, nondogmatic human beings. However, it is possible that a prevalence of unintelligent, irrational, dogmatic subjects, with demonstrated difficulties in processing probabilistic information, tend to make the human findings more conservative than the "clean" Bayesian formulas. The second hypothesis was that applying Bayesian formulas to complex, important human decisions oversimplifies the problems, and the apparent conservatism is simply a function of human intelligence responding to a recognized greater complexity. The experimental testing of these hypotheses found support for the second, but not for the first. That is, it appears the difference is not because humans are conservative but because Bayesian formulas, as simulations of human decision making, are inherently risky. The research should be of interest to psychologists and statisticians interested in Bayesian methodology. NICHHD.

#### *Personality and Test Taking Behavior.*

RAYMOND C. MULRY

To examine more carefully the relationship between test taking time and personality and situational variables the Mulry Index of Reflective Effort (MIRE) was developed. MIRE consists of "easy" and "difficult" items to which responses of agree, disagree, or undecided can be made. The easy items are defined as those with which most respondents will agree or disagree. Difficult items are defined as those to which most test takers will respond as undecided. High reflective effort is inferred when a respondent takes a great deal of time in answering easy items. Low reflective effort is inferred when respondents take the same amount

of time in answering difficult items as they do in answering easy items. Several studies have been done that deal with formal properties of MIRE. In one study, the relationship between probabilities of endorsement and item latencies—or the time it takes to respond to an item—was examined (RM-67-12). In a later study, an attempt was made to manipulate latencies on difficult items through the use of a situational manipulation. Item latencies were found to vary as a function of the probability of endorsement and as a function of motivational determinants. NICHHD.

#### *Test Anxiety Study.*

BRUCE BLOXOM

Factor analytic studies have shown that test anxiety is multidimensional—that is, it has a variety of causes rather than simply being a matter of a so-called global anxiety. For example, anxiety in test taking may come from numerous aspects that are specific to each test situation, such as the test-taker's estimate, in advance of the test, of its importance to his final grade, or his own estimate of how well he is prepared for the test. This study investigated a number of such aspects. For each aspect, the subjects who responded anxiously to it were then compared with other subjects with respect to their scores on conventional test anxiety measures and their actual performance in formal testing situations. The subjects were found to respond differently in relation to whether they felt the test to be important or unimportant in advance of taking it. On the subjectively unimportant tests some reported feeling little anxiety and performed relatively poorly. Other subjects felt anxious on both important and unimportant tests and performed neither better nor worse than the other subjects. Of the aspects studied, how well prepared a person feels he is before taking a test seems to have the greatest relation to the amount of tension felt during the test. Preparedness also interacts with a subject's typical level of performance in school and with global test anxiety in determining the tension felt during test taking. NICHHD. RB-68-30.



## Social Behavior

### *Negotiation Behavior under Conditions of Observation and Consultation.*

NATHAN KOGAN

An important question in exploring the nature of negotiation is how leaders' negotiations may be affected by the presence of their delegates, and how delegates' negotiations are affected by the presence of their leaders around or near the "conference table." This study found that delegates who observe their leaders negotiate tend to urge more risky courses of action upon the leaders than the leaders would be willing to accept for themselves. Also, leaders who observe their delegates negotiate show considerably more agreement with their delegates' decisions than was the case when the discussant and observer roles were reversed. This is quite possibly the first study of how negotiators perform under surveillance, which so very often prevails in "real world" negotiation settings. A flow chart of the complex methodology involved is available to researchers. The research should be of interest to social psychologists, sociologists, political scientists, and other social scientists concerned with negotiation and conflict resolution. ARPA. RB-69-55 (H. Lamm, G. Trommsdorff, co-author).

### *Individual and Group Decision Making under Conditions of Responsibility for Others.*

NATHAN KOGAN

In virtually all decision-making studies, the consequences of the decisions are borne only by the decision makers themselves. Yet it is common knowledge that decisions often affect others who had no voice in the decision. This study sought to determine whether there is any difference between individuals and groups in making decisions that affect others. It was found that individuals are more cautious in making decisions affecting others than they are in making decisions that affect only themselves. Groups, however, take about the same risks whether the decisions affect themselves or others.

The procedure involved a choice of wagers for self and others, with real monetary payoffs. Subjects were students at the Sorbonne. The research should be of interest to social and experimental psychologists with special interest in decision making. A noteworthy feature is that it links group risk-taking studies to research into social responsibility and helping behavior. The research seems particularly relevant to the problem of promoting the welfare of others in complex decision-making contexts. ARPA. RB-69-9 (M. Zaleska, co-author)

### *Determinants of Risk Taking for Others.*

NATHAN KOGAN, ALLAN TEGER

This study is an extension of the question explored in "Individual and Group Decision Making under Conditions of Responsibility for Others." Here the specific question was: When people make decisions for others, what will be the effects of friendship and reciprocity between the decision maker and the person for whom the decision is made? That is, what will be the effects of whether friendship is present or absent, and whether reciprocity (both parties having something to gain or lose from the action of the other) is present or absent? The study found that friendship had no significant influence on decision making in this situation. Reciprocity, however, had significant effects—subjects were more cautious in making decisions that affected others when some degree of reciprocity was involved. The research was carried out at Douglass College, with wagers with real monetary payoffs. It should be of interest to social and experimental psychologists with an interest in decision processes. NICHHD.

### *Effects of Anticipated Delegate Status on Group Risk Taking.*

NATHAN KOGAN

Will prior knowledge that one is going to be a delegate defending a group's interests affect one's tendency to take risks? Exploring this question, this

study found that on the whole anticipated delegate status neither raises nor lowers preferred risk levels in comparison with a control group. However, differences were observed between delegates who were selected randomly and those who were elected by the group. The randomly selected delegate seemed to have little influence on the group's choice of a risk level. In contrast, the chosen delegate had considerable influence on the group's choice of risk level, but sometimes of a backlash variety—that is, efforts to dominate the group can meet with rejection of the would-be dominator by the group. An aspect of interest is that the study was conducted at Nanterre in France shortly before the student riots of May and June 1968. The research is reported in *Acta Psychologica*, 1969, 29, 228-243, (W. Doise, co-author), and should be of interest to social psychologists, political scientists, sociologists, and all concerned with small group behavior. It extends the frequent finding of an intragroup risky-shift effect to the case in which group members later expect to perform as delegates. ARPA.

### *Cross-Cultural Study of Group Risk Taking.*

JULIA CARLSON

A curious phenomenon that has intrigued social psychologists during the past few years is the "risky-shift." If one gives a small group a problem of a certain type to discuss and requires that they must achieve a consensus on its solution, the risky-shift will generally occur. That is, the members of the group will start out with relatively conservative or "safe" views on the problem's solution, but group discussion generally leads to consensus that is markedly more venturesome or risky. This study was a cross-cultural replication of studies by two leading risk-taking theorists, Nathan Kogan of ETS and the New School for Social Research and Michael Wallach of Duke University. One sample was of 250 Africans, and the other sample was of 100 Amer-





ican students. The question being explored was whether the risky-shift might be, in part, a result of the value placed on taking risks as a matter of cultural emphasis in America. In support of the hypothesis, it was found that African subjects were less willing to advocate risks in the situations presented than the American subjects. Though the primary cause of the risky-shift still remains unknown, this study suggests that the values attached to risk taking are important determinants. The research should be of interest to social psychologists interested in the risky-shift, group decision making and tolerance for risk, and cross-cultural studies.

*Risk Taking in the Context of Intergroup Negotiation.*

NATHAN KOGAN

The problem of negotiating solutions to conflicts is a meeting point for psychologists, political scientists, and international relations experts. Within psychology, an important facet of research has centered on risk-taking phenomena—that is, the ways people and groups take the various risks that are involved in reaching decisions acceptable to *both* parties in a conflict. This study sought to find out whether the risky-shift effect found within one group might also be found in negotiations between groups. The main finding was that the risky-shift effect breaks down when negotiators are strongly committed to a position on an issue which their reference group has taken prior to negotiation. However, negotiators with a minimal commitment to a prior position continue to manifest risky-shift effects. The study was carried out at the University of Mannheim, and will be reported in the *Journal of Experimental Social Psychology* during 1970 (H. Lamm, co-author). The research should be of interest to social psychologists, political scientists, international relations experts, and other social scientists working on conflict resolution. ARPA.

*Pessimism-Optimism and Risk Taking.*

NATHAN KOGAN

Research has shown that groups tend to shift toward more risk taking in a

standard experimental situation. Given this fact, the question has arisen whether there might also be a shift toward optimism, in the sense of the group assigning a higher probability of success to the risky alternative. Cognitive dissonance theory would predict such a result. However, this study found that, contrary to predictions, the groups became progressively more pessimistic in the risk-taking situation. Also, the relation between risk and pessimism was found to be asymmetrical. That is, risky-shifts are followed by pessimistic shifts, but pessimistic shifts are not followed by risky-shifts. The sample was of German students at the University of Mannheim. The research is within the risk-taking studies tradition and attempts to integrate the social psychological work concerned with comparing individual and group products. It should be of interest to social and organizational psychologists, and will appear in the *Journal of Personality and Social Psychology* during 1970. (H. Lamm, G. Trommsdorf, co-authors). The results are also important in attempting to generalize risky-shift work to the real world in that they indicate that pessimism may act as a reality brake on impulsive action by groups. ARPA.

*Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma.*

NATHAN KOGAN, MARGARET HERMANN

Over the decade or so since its first proposal by Rapoport, the somewhat involved "prisoner's dilemma" situation has become the takeoff point for research in gamelike behavior and decisional processes, much of which has been applied to studies simulating international conflicts and their possible resolutions. Virtually all of the work with prisoner's dilemma, however, has involved only one person negotiating with another person, whereas in real-life conflicts there is generally some representation of others involved. This study examined the effects of individuals compared with teams in the dilemma, and also compared the effects of teams of friends with teams of strangers. Over several trials, it was

found that individuals are relatively uncooperative at the beginning of negotiations and become progressively more cooperative. Teams of friends, on the other hand, are highly cooperative at the beginning and remain high. Teams of strangers are neither high nor low in level of cooperation, but begin and remain in-between. The study also explored the possibilities of pairing "prisoners" according to personality similarities and differences, and found systematic variations in cooperating behavior according to the degree to which personalities were matched or mismatched. The research should be of interest to social and personality psychologists and other social scientists interested in games and decision processes. ARPA.

*Negotiation in Leader and Delegate Groups.*

MARGARET HERMANN, NATHAN KOGAN

From situations of the campus to those of world conflict, many hopes have come to be centered in the efficacy of negotiations and negotiators. Thus, conflict resolution is becoming an area of pressing concern in social psychology. This study explored the question of how negotiations carried out by individuals with delegated authority differ from negotiations carried out by people with leadership roles. (A real life analogue, for example, would be negotiations by relatively unknown diplomats versus negotiations by the Presidents of two countries in conflict.) A notable finding was that in negotiations leaders tend to be more influenced by other leaders than are delegates by other delegates. Another important finding was that negotiating leaders will take greater risks than negotiating delegates. This came through investigation of what is known as risky-shift phenomena—and the study is one of the first to be made of *intergroup* risky-shift. The study is reported in the *Journal of Conflict Resolution*, 1968, 12, 332-344, and should be of interest to social psychologists, sociologists, and political scientists—as well as anyone in education facing a situation of possible conflict with the necessity for negotiations to resolve it. ARPA.

### *Intercultural Attitude Study.*

HAROLD GULLIKSEN

Much as understanding individual differences concerns those interested in the advancement of the individual, understanding national differences and similarities is a concern of those interested in better international relations. A major problem is the assessment of differences in national attitudes. This study used advanced statistical methods to analyze attitude statements obtained from groups in Norway, Sweden, Finland, France, Germany, Belgium, and Poland, also Texas and Pennsylvania in the United States. People were asked for their opinions on such enduring human concerns as occupational prestige, methods of choosing a job, attitudes toward work, and aims in life. In assessing "aims," for example, it was found that "To achieve a deep and lasting love" was given top ranking by people in Poland, Sweden, Italy, and the United States. Also "To enjoy good health," "To have a happy home life," and "To bring up one's children well" were ranked next highest by Swedish, Italian, and United States groups. However, while many similarities are present, the study also found differences among the groups in the different countries. Determining the stability of individual and of group judgments is a purpose of the study. It should be of interest to those concerned with theoretical aspects of scaling and with cross-cultural comparisons.

### *Prosocial Behavior.*

DAVID ROSENHAN

Social scientists tend to view altruism as either a device to cover actual feelings of selfishness or hostility, or as something pressed into the growing child by admonitions to "be good to others." This study examined altruism through a series of depth interviews with people who risked their lives to save others, and with white people who were active in the Negro freedom movement in 1961. Interviews with those active in the freedom movement over an extended time indicated that they model their behavior after parents or other ideal figures who were also

notably altruistic and with whom they deeply identified. Among those who engaged in freedom activities over relatively short periods of time, an interesting difference came to light. They seem to have had parents who preached but did not practice altruism and were disliked by their children for this specific failing. The study is reported in two works: *Trends and Issues in Developmental Psychology* (Holt, Rinehart & Winston, 1969) and *Altruism and Helping* (Academic Press, in press). It should be of interest to child and experimental social psychologists, researchers in higher education, and those concerned with social action research.

### *Observation of Prosocial Behavior.*

DAVID ROSENHAN

The bibles of all religions testify to the fact that character development has long been a central social concern. For years it has also concerned some psychologists, but only recently as a matter for consistent, rigorous experiments. Current interest centers on the ways by which good or bad models (parents, teachers, even television heroes) may influence a child's character development. In a study of how altruism is encouraged, two groups of children participated in a bowling game, one group with an adult model present and one group without. Each time the model won he received two rewards, one of which he contributed to a hypothetical charity. Each child who won was then observed to see if he would, on his own, contribute to charity after the model left the room. Even though the children were told they might contribute, all of those who had no model to observe failed to do so. Those who did have a model contributed—but only if they had already donated before the model left the room. Later work with this experimental situation varied the age and sex of the model, the nature of the reward, and teaching conditions. RB-66-50; *Journal of Personality and Social Psychology*, 1967; and *Trends and Issues in Developmental Psychology*. NICHD and NIMH.

### *Cognitive Styles and Attitudinal Consistency.*

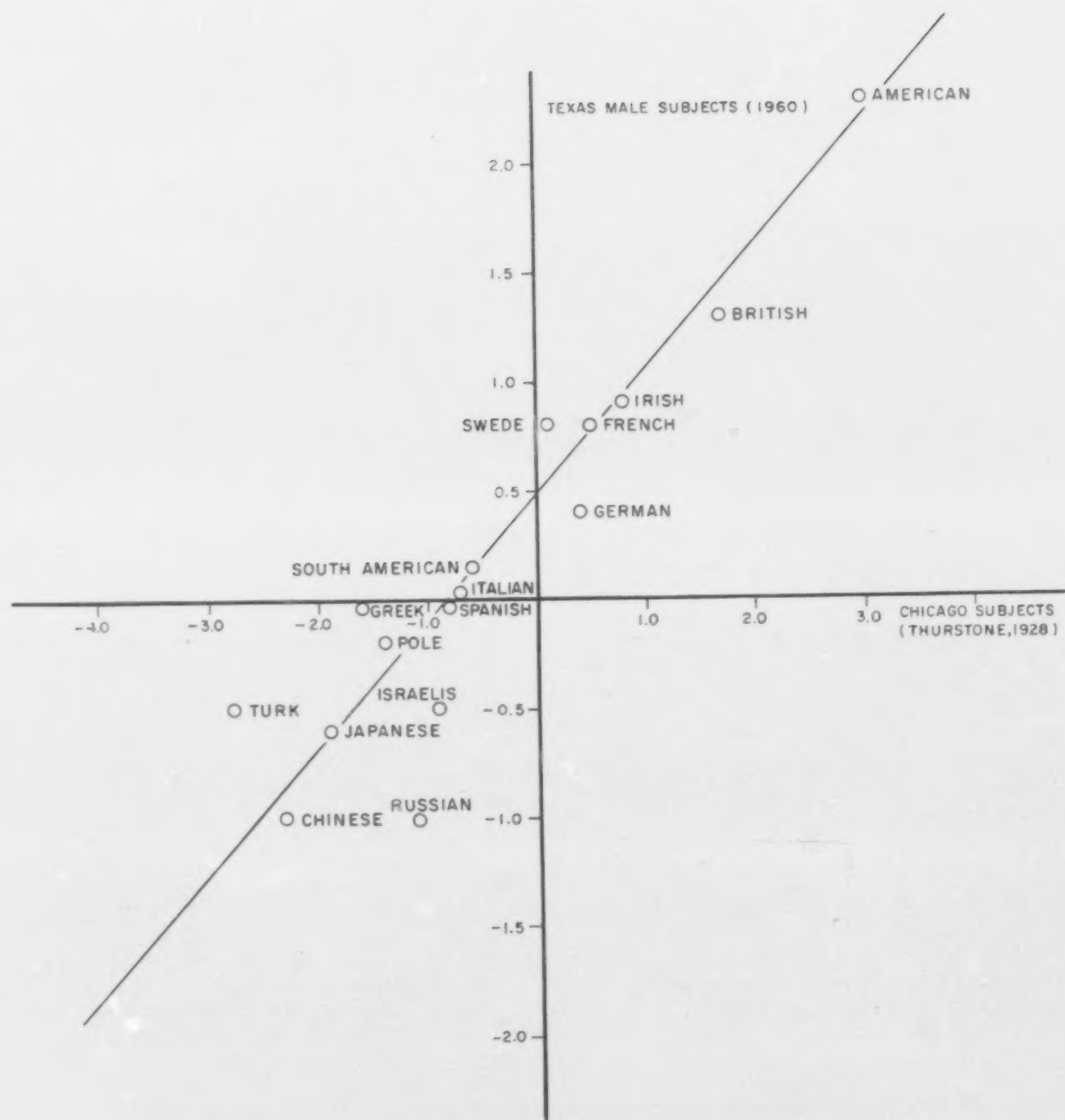
SAMUEL MESSICK

One of the major areas explored by social psychologists during recent years is the dynamics of attitudinal consistency. On the everyday level, this consistency is often quite apparent—people's beliefs being linked in often very predictable ways with such matters as their social status, economic level, occupation, political party, race, and age. Going beneath the surfaces to get at exactly *how* these consistencies work involves some very intricate experimentation and theories, but social psychologists have been spurred to the effort because of the vast range of social phenomena such work is beginning to clarify. A great deal of recent work, in particular, has dealt with the extent to which both cognitive and affective components of attitude systems are organized in consistent patterns, with the operation of motivational pressures toward consistency, and with the specific mechanisms whereby inconsistency is resolved. This study is exploring the relation of individual differences to attitudinal consistency and the generality of such operations across different measurement approaches and response changes. A notable methodological aspect of the study is that the attitudinal consistency of a sample of 200 high school seniors—who have been extensively tested with other measures—is being measured using the experimental procedures of four major studies in this area: the McGuire Logical Consistency Test, the Morrisette Balance Completion Test, the Osgood and Tannenbaum Congruity Test, and the Rosenberg and Abelson Fenwick Test. The study is being done in collaboration with J. E. Singer of the State University of New York at Stony Brook. When completed, the research should be of interest to social and personality psychologists working in the area of attitudinal consistency and cognitive styles. NICHD.

### *Dimensions of Conformity.*

LAWRENCE J. STRICKER,  
SAMUEL MESSICK, DOUGLAS N. JACKSON

Research into the processes involved in  
(continued on page 51)



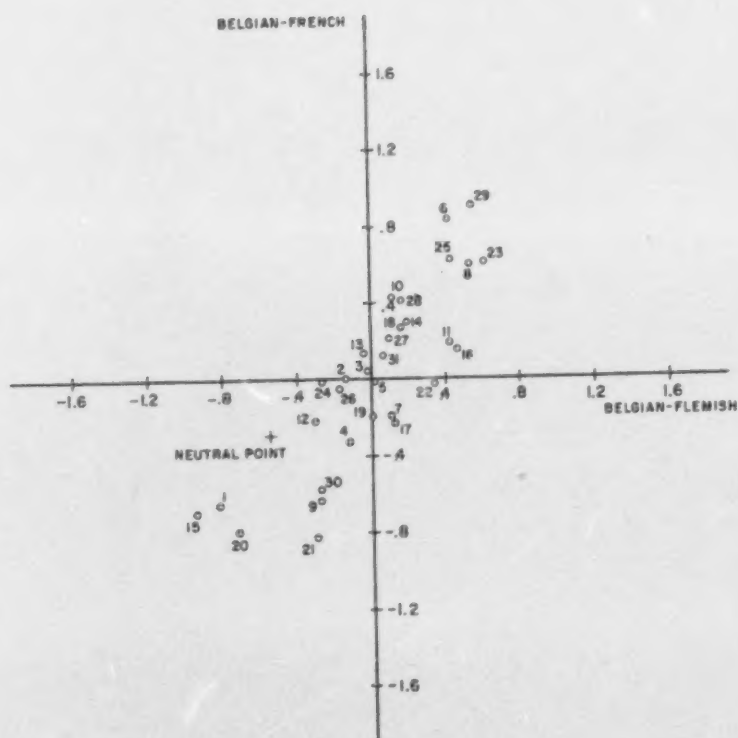
COMPARISON OF SCALE VALUES FOR NATIONALITY PREFERENCES. (1928 vs. 1960)

How people of one culture perceive people of other cultures is a question of many facets for the measurement specialist. The national preferences of some Texas college students in 1960 are shown in this chart, as compared with the national preferences of students in Chicago in 1928. In both years and both localities students rated Americans as their first preference among nations (not unsurprisingly, the study found that all people surveyed preferred their own nation first). At the low

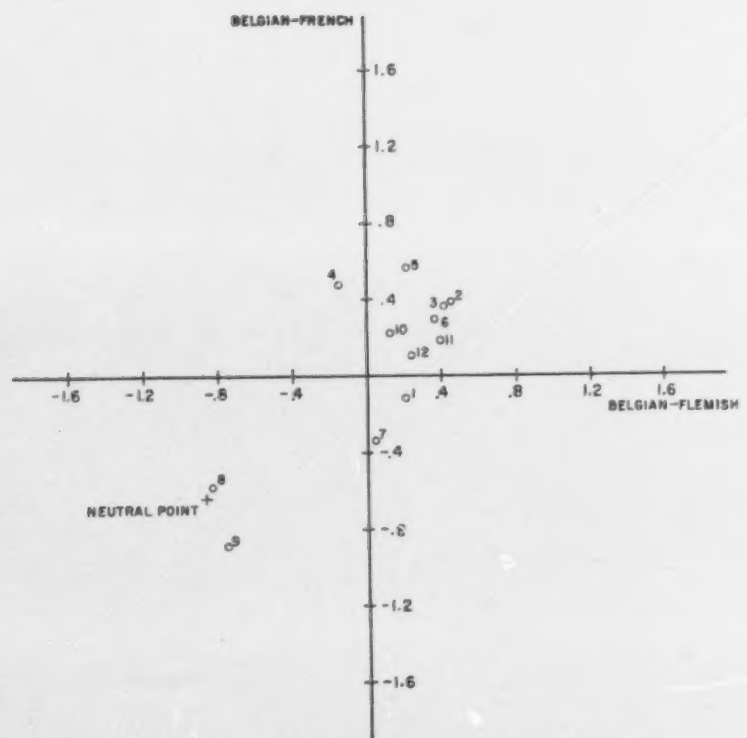
end of the preferences the greatest changes were shown in attitudes toward Turks and Russians. While rated lowest in national preference by the Chicago students in 1928, by 1960 Turks had moved to fourth from the lowest in the preferences of Texas students. Meanwhile Russians had gone from sixth lowest in 1928 to lowest in national preferences for Texas students by 1960.

(See Gulliksen, *Intercultural Attitude Study*.)





COMPARISON OF SCALE VALUES OF REASONS FOR WORK:  
BELGIAN-FRENCH AND BELGIAN-FLEMISH.

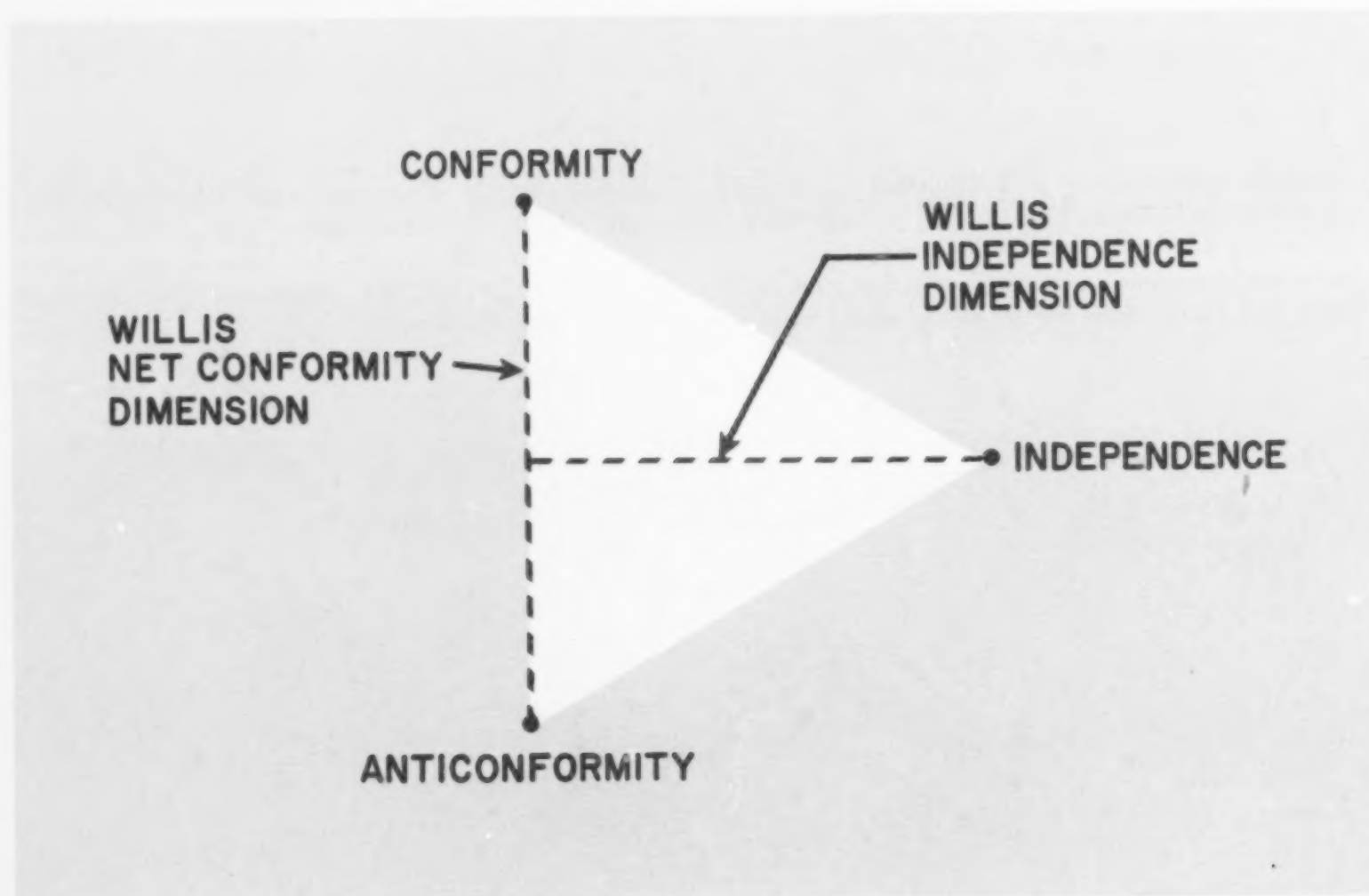


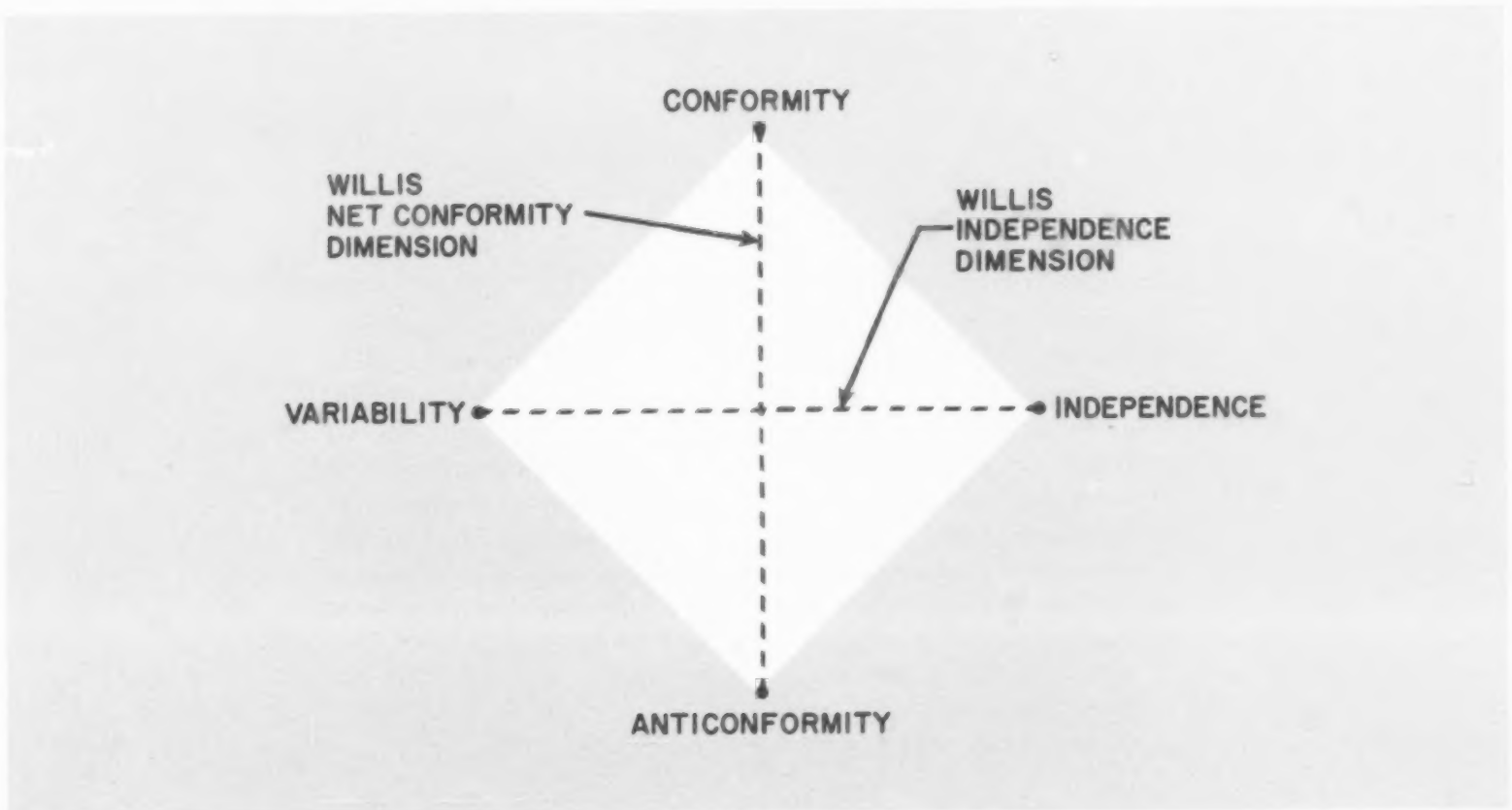
COMPARISON OF SCALE VALUES OF GOALS OF LIFE:  
BELGIAN-FRENCH AND BELGIAN-FLEMISH.

Measuring the degree to which the attitudes of people of various cultures differ from or are alike one another is a pursuit with practical applications to fields ranging from education to diplomacy. A questionnaire was given to people in Belgium, Italy, France, Germany, and the United States. These charts indicate the degree to which Belgian-French and Belgian-Flemish agreed or disagreed on goals of life and reasons for work. The charts indicate, for example, that while both types of Belgians were in close agreement on "doing God's will" as a goal of life, (point 2, in quadrant II), they greatly disagreed on the importance of "self sacrifice for the sake of a better world" as a goal (point 4, quadrant I). Likewise, in rating reasons for work they agreed on the importance of good working conditions (point 8, quadrant II), while disagreeing to some degree on the importance of job security (point 17, quadrant IV). (See Gulliksen, *Intercultural Attitude Study*.)



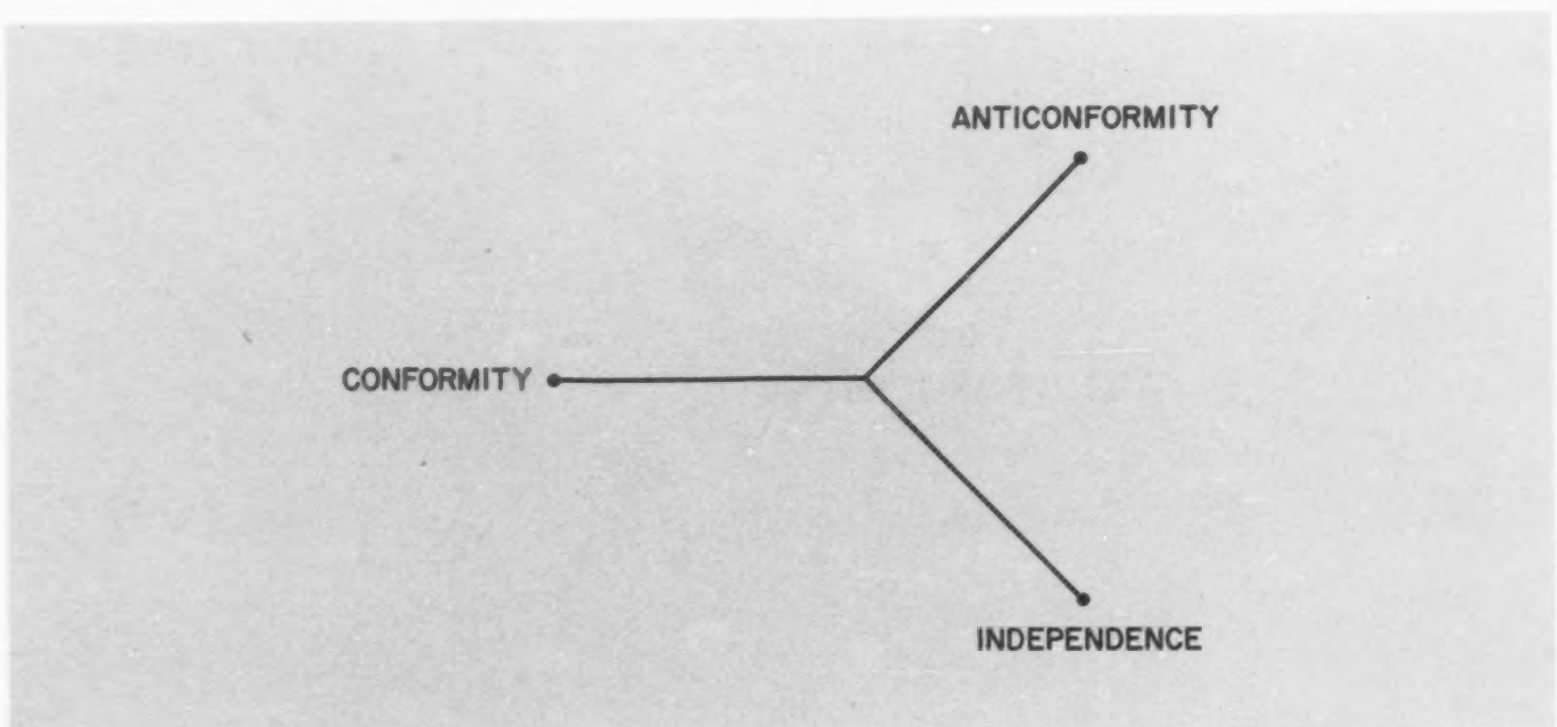
During recent years the study of conformity has centered upon perfecting models of theoretical relationships, which are based on group pressure experiments. The popular view is of conformists and non-conformists. This figure, however, shows a "triangle" model proposed by Krech, Crutchfield, and Ballachey combined with another model proposed by Willis. It indicates *two* kinds of nonconformists—an anticonformist, who is actually highly conforming in his rejection of pressures to conform; and the independent person, whose behavior is relatively unaffected by group influences. (See Stricker, *Dimensions of Conformity*.)





The above figure is of the Willis "diamond," a theoretical model that adds another interesting component—the variable person, who seems to behave independently, but actually pursues an erratic personal strategy. The figure below depicts the results of an extensive factor

analysis study by ETS psychologists Stricker, Messick, and Jackson. This result differs from both previous models, but indicates somewhat more support for the "triangle" than for the "diamond." (See Stricker, *Dimensions of Conformity*.)



social influence has raised a number of important questions. This study explored two of these questions. First, how are conforming, anticonforming, and independent behaviors inter-related? Secondly, to what extent do subjects' suspicions distort the results in research of this kind? (Standard procedures call for a deception wherein the subject is misled about both the purpose of the study and the behavior of the other subjects.) This investigation found that many subjects correctly suspected they were being deceived and did not conform; suspicion, in effect, invalidated the experimental results for these subjects. Contrary to most theoretical views, it also found that conformity and independence seem to represent one bipolar dimension, while conformity and anticonformity appear to reflect another bipolar dimension. In addition, the conformity, anticonformity, and independence behaviors observed in one experimental situation were not highly related to the corresponding behavior in other situations. This result suggests that social influence has rather specific effects, depending on the particular situation involved. This study is of potential interest to researchers in personality and social psychology because of its unusual methodological features. Among its implications are that the use of deception in psychological research may be practically, as well as ethically, unwise. Various phases of this study are reported in the *Journal of Personality and Social Psychology* (1967, 5, 379-389, and in press), and the *Journal of Experimental Research in Personality* (1968, 3, 151-167). NICHHD.

*Exploratory Research on the Escalation of Conflict.*

ALLAN TEGER

Social psychologists have advanced the study of human conflict to the point where a body of formal theory on the

subject is becoming available. Conflict theory to date has been concerned mainly with the total amount, or the presence or absence, of conflict in varying circumstances. A theory of escalation, or how conflicts grow, has been lacking. This exploratory research is attempting to develop such a theory through a series of laboratory and field studies. Variables being studied include the rate of escalation, the value of victories, changes in the scope of the conflict, changes in the images of the parties to the conflict, and changes of issues as the conflict progresses. One field study, for example, was of the occupation of a building by radicals on the Princeton University campus. So far, pilot data indicate that conflicts do not escalate at a uniform rate. Rather, the conflict may escalate more rapidly immediately after the parties to it pass a point of no return, and it may de-escalate somewhat when it passes the point at which one party had anticipated a victory. Results of the Princeton building seizure indicate that the occupation of a building will reduce student support for radical demands, while increasing support for building occupation as a tactic. The research should be of interest to social psychologists and others studying conflict mechanisms.

*Conceptual Complexity and Normative Social Expectations.*

ROY E. SHORE

Research on the development of social norms has typically dealt with their content, height and strength. (Norm components can generally be plotted as a curve on a graph, with the curve indicating height and strength.) However, another important way norms probably change in the course of development is in complexity. One question for research concerns the flexibility of norms. Are they applied differently to different people and situations? Are they based on many rather than a few considerations? Using a

theory of conceptual complexity, this study attempted to predict the nature of a change in the complexity of norms. Also using a measure based on this theory, it tried to predict an individual's normative complexity. One finding was that those predicted to have high normative complexity did in fact make more complex normative judgments. It was also found that they tended to base their norms on mutual or two-way relationships rather than authoritarian, or one-way, relationships. A noteworthy aspect of the study was the use of a questionnaire designed to reveal the structure of social expectation norms developed by Dr. Walter Emmerich of ETS. The research should be of interest to social psychologists and others studying social norms and role theory. NICHHD.

*Behavioristic Theory of Attitude Development.*

STANLEY M. ZDEP

Most researchers agree that attitudes are learned. This implicates the concept of reinforcement—or an event that strengthens a response in a particular situation. However, a single reinforcing event can have reward value for one person and be seen as punishment by another. Additionally, there has been considerable controversy over questions of which attitudes are being reinforced in situations of some complexity, as well as what determines numerous inconsistencies between expressed attitudes and observable behavior. This study reviewed existing theory and research data in an effort to define (1) how reinforcement contributes to attitude formation, and (2) what constitutes reinforcement for different kinds of individuals. The study was theoretical rather than empirical, and proposes an integrated framework for previously diverse elements of research and theory. It should be of interest to educators interested in the formation of negative attitudes toward school among some groups in grades 1-3.

## II

---

# EDUCATIONAL APPLICATIONS



## Early and Primary Education

### *Assessment and Development of Cognitive Skills in Young Children.*

EDWARD A. CHITTENDEN, GERALD W. BRACEY, ANNE M. BUSSIS, MASAKO N. TANAKA, and JEAN H. OROST

Known originally as *The New York City Project*, this venture has developed and field-tested new teaching-testing materials in an area of crucial need. The project started as an attempt to develop assessment methods that would be particularly appropriate for urban children. As the project progressed, emphasis solely on assessment shifted to include teaching as well as testing. In essence, the child-development theories of Jean Piaget have been translated into a classroom approach that combines teaching with testing in a new way. The approach, now in use in many schools throughout the country, was originally developed for use with first graders. Current work is directed toward producing a set of materials for teachers of 4- and 5-year-old school children. Beyond assessment *per se*, an important research emphasis has been to see to what extent the approach can be used to encourage cognitive growth. Studies of an instructional nature were undertaken during 1967-1968 in the areas of conservation of number and classification concepts. The project should be of interest to teachers and school administrators and to researchers in early childhood. First grade materials and other information on the project are available through the Cooperative Tests and Services office of ETS. The Fund for the Advancement of Education and the Carnegie Corporation have been the major funding sources. RM-68-13. RB-68-49.

### *Preschool Television for Children.*

SAMUEL BALL

By now the new experimental television show *Sesame Street* has become a welcome fixture in many American homes. The original idea was to use the creative experience and ingenuity of commercial television people (artists, writers, film makers, etc.) not just to

sell soap or package "normal" entertainments, but to see if the same methods might be used to pipe education directly to young children in their homes. One device, for example, has been the use of the one-minute commercial to inject the ABC's or number concepts into entertaining programs offering more indirect learning values. From July through December of 1968 consulting psychologists from ETS worked with Children's Television Workshop producers to establish behavioral objectives—that is, the specific kinds of learning that might best prepare preschool children for a good school experience. They also worked to develop procedures that would measure a child's progress toward these objectives, and measure how much the program seemed to be helping. Then on the basis of this experience, from January through June of 1969, the ETS group developed the necessary tests, and in July the new television teaching approach was given a trial run with 60 children who were carefully pretested and posttested. Variables being measured included knowledge of letters, numbers, forms, body parts and relational terms, and skill at classifying and sorting. An approach for interviewing parents was also developed and tried out during this time. The research should be of interest to preschool educators and researchers in educational technology. A report will appear during 1970 in *Educational Television*.

### *School Readiness Criteria Study.*

SCARVIA B. ANDERSON

In both the popular and professional literature on early schooling the high percentage of "unready" first graders has been noted. This has raised the question whether our customary first grade education is relevant to large numbers of first graders. It also underlies the concern of educators for achieving the right emphases for Head Start in comparison with other preschool and first grade programs. This study sought to find out what "school

readiness" means to teachers and child development experts, and to determine the relation of teachers' and experts' readiness criteria to the generally stated objectives of preschool education. Some 7,000 behavioral indices (can tie his shoes, can take directions, etc.) were collected from a national sample of 300 first grade teachers. Comparison criteria and objectives were developed through panel discussions of experts. One finding was that teachers are likely to define "readiness" in cognitive terms but "non-readiness" in personal-social terms. There was also some discrepancy between teacher and expert expectations for beginning first graders and the stated objectives of preschool education programs. The research should be of interest to educators concerned with preschool and elementary programs, test makers, and researchers concerned with criterion definition.

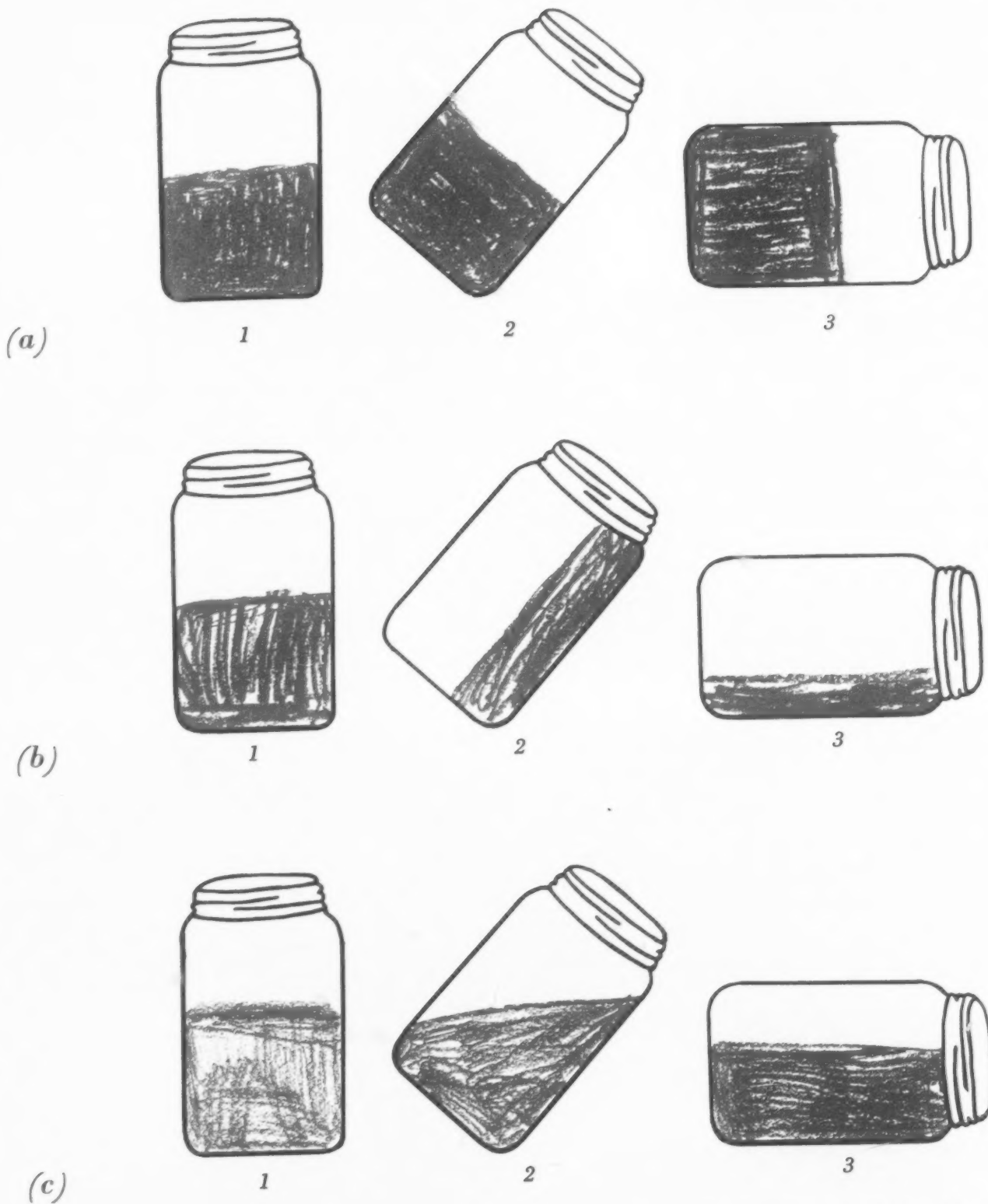
### *SEARCH.*

FREDERICK R. KLING

Automated instruction for preschool children presents both unusual opportunities and unusual problems. Young children are curious, eager to learn, sometimes even desperate for someone to read to them. They also have short attention spans, communication problems, and minds of their own. Finding an effective combination of equipment and programming ideas to maintain the interest of prereaders and to stimulate their learning is the goal of project SEARCH (Sequential Enrichment and Reading for Children). Doing this at a practical level, to obtain some of the advantages of computerized equipment without computer costs, has been a further objective. A prototype teaching machine and illustrative audiovisual programs have been developed. The machine asks a question and the child answers by selecting a portion of the picture and pressing one of three buttons. Depending on the child's answer, the machine then branches onto different

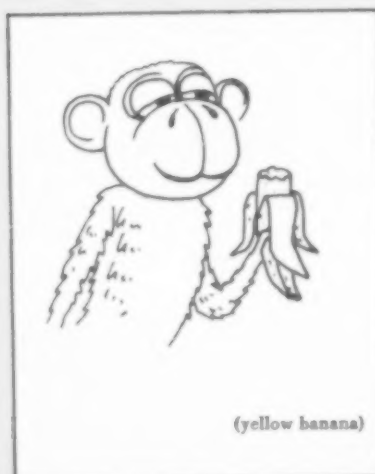
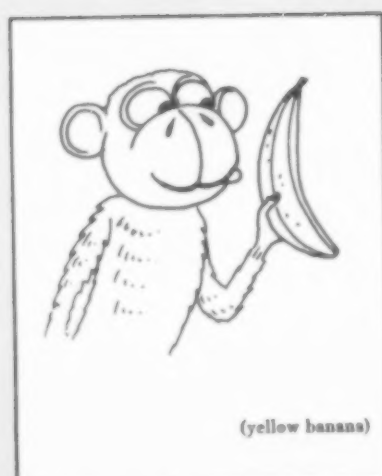
(continued on page 58)

### Water Level Task



An important cognitive skill involves a child's understanding of relationships within our physical surroundings. The "water level task" shows how children learn that vertical and horizontal are always in relation to the earth's surface. These actual classroom responses show the range from (a) preoperational, through (b) transitional, to (c) concrete operational thinking. (See Chittenden, *Assessment and Development of Cognitive Skills in Young Children*.)

Set of Cards for the Sequence Completion Task

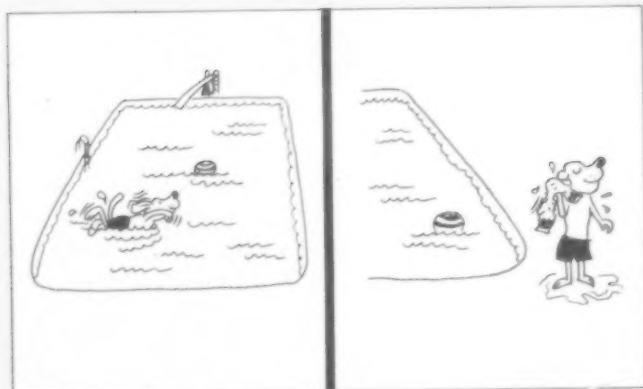


Many studies indicate that a child's awareness of the concept of time undergoes great changes in the early years. Thus, the development of the concept of time can be a useful indicator of cognitive growth in young children. These are the cards for a simple *sequence completion* task. Given the two cards shown above the heavy line, the child is asked to complete the sequence using one of the three cards shown below the line. (See Chittenden, *Assessment and Development of Cognitive Skills in Young Children*.)



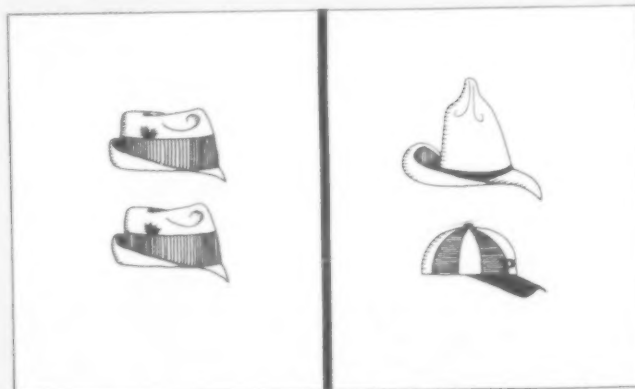
COPYRIGHT © 1967 BY EDUCATIONAL TESTING SERVICE, ALL RIGHTS RESERVED.

*Sample Items from the Language Comprehension Task*



*Dog is swimming*

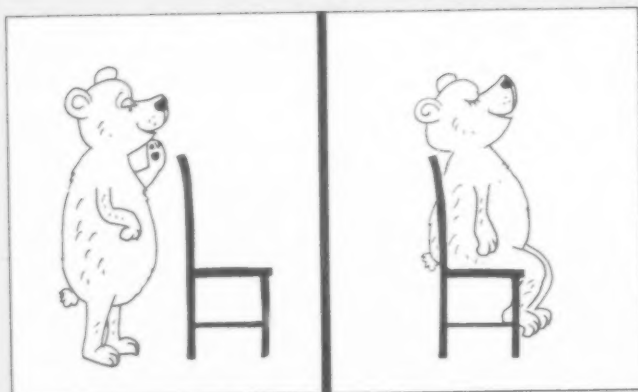
*Dog swam*



*Same hats*

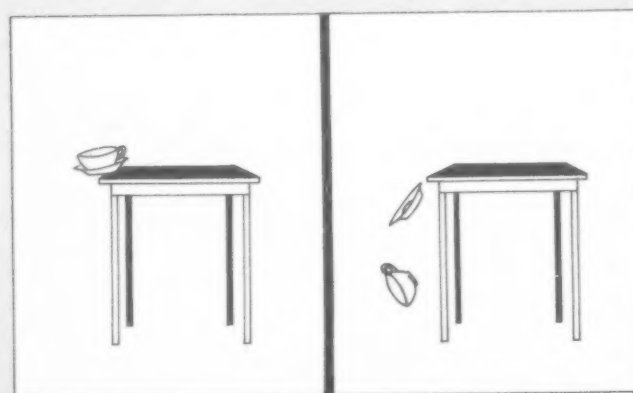
*Different hats*

When do children begin to comprehend the small words and inflections that govern syntactic structure and logical meaning? These sets of matched pictures from the ETS *Language Comprehension Task* are used in research designed to answer such a question. (See Chittenden, *Assessment and Development of Cognitive Skills in Young Children*.)



*Bear is not sitting*

*Bear is sitting*



*Cup will fall*

*Cup is falling*



teaching tracks, thereby individualizing the instruction. The immediate response of a human voice to the child's answers appears to be of crucial importance to the program's effectiveness. The research should be of interest to teachers and educational technologists.

*Literature Appreciation Test,  
Grades 4-6.*

RODERICK IRONSIDE

To devise a test that would measure literature appreciation might, on the surface, seem rather easy to do. Some who have tackled the task in the past, however, have questioned its feasibility. This project has developed a measure that not only will assess sensitivity to aspects of literature (such as literary effects, authors' devices, and styles and meanings) but also will measure *changes* in sensitivity over a period of time—an essential capacity for an instrument intended to measure the effects on students of teaching, reading, and other exposures to literature and life. The completed test, "A Look at Literature," and its handbook are now available through Cooperative Tests and Services at ETS. The test can be used by teachers and researchers to assess the effects of various teaching approaches on students, to determine growth in appreciation over several years, or to enhance literature instruction itself. The measure is intended for group or classroom, rather than individual, comparisons. "A Look at Literature" was developed in cooperation with the Research Foundation of the National Council of Teachers of English.

*Church School Curriculum:  
Readability and Usability.*

RODERICK IRONSIDE

In many churches Sunday school teaching is a form of amiable chaos at best. Fairly recently some large denominations have tried to improve the situation with modernized materials and teaching methods. This study sought to determine the actual readability and usability of some of the new materials. Over 100 samples of Sunday school prose were analyzed in terms of various readability formulas. Then 25 percent of these samples were tested with 600 subjects using the relatively new CLOZE method of judging difficulty of material in terms of student reading behavior and the reactions and test answers of students and teachers after the material had actually been taught in church school settings. It was concluded that much of the material studied was too difficult for the grade levels for which it was intended. It was also found that readability formulas alone are not sufficient for establishing readability and usability of materials. The study indicates the necessity of directly testing the materials with readers of the type for whom the materials are ostensibly intended. The point is of importance as most assessments of this kind over the past 10 years have been made with readability formulas and relatively little work has involved more direct methods—and virtually no studies have been made of such church-related curricular materials. The study should be of interest to those concerned with church education, assessing the difficulty and acceptability of text materials, or validating readability formulas.

*A Study of the Interrelationships Among  
Reading Comprehension Ability,  
Writing Ability, and the Linguistic  
Structure of the Written Language of  
Sixth-Grade Children.*

CAROLYN E. MASSAD, JOSEPH P. KENDER

There is a new surge of interest in how the four language functions of listening, speaking, reading, and writing develop in children. It is presumed that the four abilities are interrelated; however, the interrelationship between reading and writing has been relatively overlooked in research. This study sought to find out whether a significant relationship exists between the written language ability and the reading comprehension ability of sixth-grade pupils. One analysis was made of the linguistic structure of children's freely written language similar to the analysis used by Walter Loban in studying children's oral language. The study found that two aspects of the children's freely written language—namely the number of communication units and the number of words per communication unit—are not significantly related to their reading comprehension ability, nor to their ability to use the conventions of language to express themselves. Further research on the linguistic structure of children's language as expressed in the free essay, for comparison with data from objective measures of reading and writing, should shed more light on the problem and provide information helpful in developing guidelines for instruction in reading and writing. The study should be of interest to researchers in psycholinguistics, language development, reading, and language arts, and to educators.

## Secondary Education

*The College Board Candidate Survey.*

ELIZABETH W. HAVEN

Probably more changes have occurred in high school curricula over the past 10 years than in any previous decade in our history. To gain a comprehen-

sive picture of what achievement test candidates were actually studying by the mid-sixties, the College Entrance Examination Board supported a survey of about 38,000 students who took College Board Achievement Tests in more than 7,500 high schools through-

out the United States during the 1965-66 academic year. Replies were received from 26,000 students. On the basis of this high return (70 percent) a great deal was learned both of changes and what remained unchanged in the major subjects being taught in

American high schools. Test Development Reports written by ETS subject matter specialists that summarize findings in the following areas are now available from ETS at 50 cents per copy: Latin (TDR-69-1), Modern Foreign Languages (TDR-69-2), Physics (TDR-69-3), Mathematics (TDR-69-4), Chemistry (TDR-69-5), Biology (TDR-69-6), History and Social Studies (TDR-69-7), and English (TDR-69-8). A sample finding in social studies was that by mid-senior year, despite the stress of curricular reformers on the need to add or strengthen courses in non-western civilizations, 28 percent of the students reported no work in history of Africa south of the Sahara, and substantial percentages said they spent less than two weeks on areas such as Latin America, Russia, Asia, and North Africa. In English, contemporary authors seemed to be gaining in popularity on reading lists—Ogden Nash was as widely read as William Blake, and J. D. Salinger was as well known as Sir Walter Scott or Jane Austen. The findings should be of special interest to teachers in the above subject areas, and to school administrators, curriculum developers, test specialists, and college admissions officers.

#### *College Board Activity and Participation Statistics.*

BARBARA PITCHER

Every year statistics on the numbers and abilities of candidates taking the Scholastic Aptitude Test (SAT) are compiled for planning purposes. These statistics can, for example, be related to population figures such as numbers of children born a number of years earlier, numbers of high school graduates, numbers of students enrolled for the first time in colleges, and sometimes important trends are revealed. This study prepared descriptive statistics on candidates during the 1967-68 testing year. It found that the number of students taking the SAT is continuing to increase, and at a higher rate than in the past two years. For example, there was an 8 percent increase in the number of SATs administered during 1967-68, compared with an increase of 3 percent during the year before, and only 1 percent during the year before that. It also found that the

increase was greater at the eleventh grade than at the twelfth grade level—about 18 percent for grade 11 and about 4 percent for grade 12. The research should be of interest to administrators and planners in the College Board, to ETS administrators, and to planners in secondary and higher education. It may also be of interest to researchers in secondary and higher education and test development. CEEB.

#### *Curricular Appropriateness of the College Board Science Achievement Tests.*

RAYMOND E. THOMPSON

One consequence of the relatively rapid change in curricula over recent years has been the need to fit testing to this change. However, where curricular change can sometimes be accomplished relatively quickly (within a single school or system, for example) test developers face a problem of another sort and magnitude. Offering tests that must be standardized and generalized to fit the needs of many systems, they must include questions relating both to "old" as well as "new" curricula, and hopefully as many questions as possible that are equally good for both. The development of several new science courses during the past decade has raised the question of how well the College Board Science Achievement Tests have met this problem. This study sought to find out whether the tests are equally appropriate for students of both the traditional and the newly developed science courses. Teachers of both traditional and new courses rated the appropriateness of test questions. Questions that were rated equally appropriate for students of both courses were considered to be course-free measures of achievement. When mean scores on the complete tests were adjusted to take account of performance on the course-free measures of achievement and on the Scholastic Aptitude Test, the results indicated that recent tests are indeed appropriate for students of both traditional and new courses. The research should be of interest to high school science teachers, science educators involved with developing new courses, educational researchers, and the general public. CEEB.

#### *U. S. National Testing in French as a Foreign Language.*

JOHN B. CARROLL

Tests of proficiency in French will be given to a large sample of school children in 1971 to compare achievement of U.S. students with students in other lands, and to find out what influences achievement in language by considering variations among nations. A final revision of the tests has been completed. During 1970 and 1971 these tests will be administered as part of the International Study of Educational Achievement, a project now based in Stockholm, Sweden. There have been few such international studies of educational achievement with tests uniformly given across nations. The final results should be of interest to foreign language teachers, educational researchers, and to educators in general. The tests relate to those in other subject matter areas—in 1967, for example, T. Husen published *An International Study of Achievement in Mathematics*. Research such as this should lead to a better understanding of the influences on school learning, particularly in the learning of foreign languages.

#### *Tests of Cultural Development.*

PAUL B. DIEDERICH

Above age 13, or the eighth grade level, average and below-average readers report less and less independent reading, and students who believe that they have no talent drop out of music and the visual arts. Academic and social pressures contribute, but for what other reasons might this break in cultural development occur at puberty? For example, are there particular difficulties in adult books that discourage the less capable readers? Tests on 20 literary works that are most widely taught in grades 9-12 are being prepared for Cooperative Tests and Services at ETS. The editing, tryout, and statistical analyses of many more questions than will be needed in the final forms of these tests is making possible an exploration in depth of possible adult book difficulties. When completed, the research should be of interest to teachers and educational researchers.



### *Newspaper Tests*

PAUL B. DIEDERICH

The American Newspaper Publishers Association Foundation sponsors a "Newspaper in the Classroom" program that reaches over three million students in grades 7 through 12 with 68 million copies of newspapers every year. The Foundation sought ETS's help in developing tests to measure some of the effects of this program. Tryout forms of tests based on two simulated newspapers were given to 4,000 students in 61 schools. Half the students were in newspaper classes and the other half in control classes of equal reading ability but without newspaper training. Significant differences were found between the newspaper and control classes, and the types of competence revealed by the test items were classified into 30 categories that correspond to most of the objectives of newspaper programs. The tests have been refined and are now available from Cooperative Tests and Services at ETS. The research should be of interest to teachers of social studies and test developers.

### *Foreign Language Innovative Curriculum Study.*

THOMAS S. BARROWS

The Foreign Language Innovative Curriculum Study (FLICS) is a project to develop new ways of teaching and testing knowledge of foreign languages. Educational Testing Service has developed tests for two different approaches to teaching French humanities. One approach is the learning laboratory method, in which in addition to his regular classroom work the student listens to tape-recorded lessons in French in a private cubicle (or carrel, as it is known). The other approach is team teaching. Teachers from various disciplines give lectures relating their specialty to French—for example, a music teacher will tell of French music, a historian of French history, and so on. Besides developing tests for both approaches, ETS provides consultation on other phases of evaluation of the project, which is being carried out by the Ann Arbor, Michigan, Public School System with U.S. Office of Education Title III support. The study

should be of interest to teachers of foreign languages.

### *Brevard County Project.*

DONALD A. TRISMEN

In the last three years of high school, most students find themselves in a quandary at the start of each school year. In general their question is: What choice of classes will give me the best grounding and most useful marks in relation to some reasonable career or college goals? The problem can be exceedingly complex and is compounded by the feeling that one is locked in with one's choices for the better part of a year. This project is an experiment designed to give tenth-through twelfth-year students in an ungraded high school in Melbourne, Florida, the information they need to plot their own courses initially, and also to change them as needed. At the beginning of the ninth year, each student takes an achievement test battery. At the beginning of the tenth year, the results of this testing are used as the basis for predicting what he may expect as his year-end achievement according to various levels of difficulty for each subject he wishes to take. The student uses these predictions to plot his course of study for the year. Then every quarter thereafter he is tested again and can replot his course. He may at this time, for example, advance himself a level of difficulty in a subject in which he is doing well and wishes to explore in greater depth, while dropping a level in another subject about which he may have been overconfident. This research into the development of curriculum-tailored progress tests, and combination of discriminant analysis and multiple-regression techniques, should be of interest to school administrators, guidance counselors, and teachers. USOE.

### *Students' Evaluation of the Advanced Placement Program and College Policy.*

PATRICIA L. CASSERLY

The College Board's Advanced Placement Program (AP) makes it possible for secondary school students who excel in any of 11 subject areas to take college-level work in secondary school.

Colleges then may or may not allow AP students to bypass the usual freshman course required in their AP subject and move on to more advanced work. Since allowing credit is up to the colleges, many have wondered whether they were being too conservative or too lenient in their placement of AP students. This research was designed to answer this and other questions through in-depth interviews with Advanced Placement Program "graduates." Three hundred and fifty-eight such students at 20 highly selective colleges were asked to evaluate their advanced preparation in secondary school, their reception and placement at college, and the effects of this experience on their overall college experience, on their plans for further education, and on their choice of a career. Eighty percent of those students who were placed ahead by the colleges felt that the Advanced Placement Program had prepared them well for work in college. These students also won honor grades—that is, A's or B's in the next sequent college course. However, half of the AP students interviewed had not been granted placement by their colleges, and of this group a majority expressed frustration and boredom with first-year courses. To them these courses seemed mainly repeats at a lower level of work they had already covered in secondary school. One consequence was that many of these students declined to take further courses in fields they had originally excelled in. The research should be of interest to teachers and administrators both on the college level and in secondary schools, and to administrators of the College Board Advanced Placement Program. CEEB, RDR-67-68, No. 5. RM-68-12.

### *The Exeter Alumni Survey.*

RICHARD E. PETERSON, BRUCE K. ECKLAND

Phillips Exeter Academy wanted to assess its impact on its students over past years in order to plan its future—a problem of obvious importance to many schools. Two special questionnaires were developed after many interviews and were sent to 13,000 Exeter alumni and to present students

and faculty. In addition to gathering extensive information on the life histories of the alumni, the forms asked for their feelings, thoughts, doubts, and convictions about the Exeter experience and its effect upon their lives. A comprehensive statistical report was made to Exeter. These data may be the basis of interpretive studies by the school. Among aspects of a more general interest are what the data may reveal of the strengths and weaknesses of the Exeter type of approach—that is, the education of an intellectually elite group within the situation of maximum control of the learning environment. Theoretically, the boarding or residential school provides such an environment in contrast to schools attended only a few hours daily while the student lives at home. PR-69-9. Also a series of articles in the Exeter alumni magazine by Bruce Eckland.

*A Study of Academic Prediction and Growth in Grades 5 through 12.*

THOMAS HILTON, WILLIAM GODWIN, CHARLES WERTS, CATHLEEN PATRICK, STEPHEN PENSAK, INGEBORG STIEBRITZ

For a long time educators have felt that extensive longitudinal studies might eliminate some of the guesswork underlying many areas of American education. Major questions of curriculum, teaching methods, and academic prediction, for example, have seemed to hinge on more thoroughly researching the particular interaction of a child with his environment that accounts for intellectual growth. This project, better known as The Growth Study, was launched for such a purpose in 1961. Thirty-four thousand students in 173 U. S. elementary and secondary schools were given a battery of tests covering a wide range of abilities and requiring about 10 hours of testing time. This same group of students was then re-tested in 1963, 1965, 1967, and finally in the spring of 1969. Tests used over this eight-year span have included the School and College Ability Tests (SCAT), the Sequential Tests of Educational Progress (STEP), a Test of General Information (TGI), College Board Achievement Tests in American History and English Composition, and the Preliminary Scholastic Aptitude Test (PSAT). The giant resultant "data bank"

has been used for a number of studies reported elsewhere in this and earlier summaries of ETS research.

*Growth as a Function of the Amount of Training.*

THOMAS L. HILTON, GOSTA W. BERGLUND

Growth study data were used in this project to determine relationships between the amount of training in math that a student receives and his actual performance. Results of two tests, STEP Mathematics and SCAT Quantitative, furnished the trend lines for math performance by students within the entire Growth Study sample. These results were then compared with the amounts of student exposure to math courses as reported in the Background and Experience Questionnaire (BEQ). With one exception, the net effect was as predicted—that amount of training is reflected in gains in performance. It was also found that this gain is discernible as early as the fifth grade. The one exception was the failure of the group reporting the largest number of math courses taken to show the gains that might be expected from such an exposure. The reason for this may be that many students reporting "high exposure" are exaggerating. The hypothesis is being checked. The research should be of interest to educational researchers.

*PSAT-AIM Norming and Follow-Up Study.*

MARTIN R. KATZ, LILA NORRIS

The Preliminary Scholastic Aptitude Test (PSAT) is widely administered to high school students to help them gauge their verbal and mathematical abilities. The assessment of these abilities is often useful in making plans for post-secondary education. Norms for the PSAT enable the student to compare his own verbal and mathematical abilities with those of various reference groups—for example, with all the students in the same grade throughout the United States, or with students who later go on to accredited four-year colleges, or with students who later go on to two-year accredited colleges. This study first obtained PSAT norms for national samples of tenth, eleventh, and twelfth graders who were tested in the fall of 1966. The eleventh grade sample was

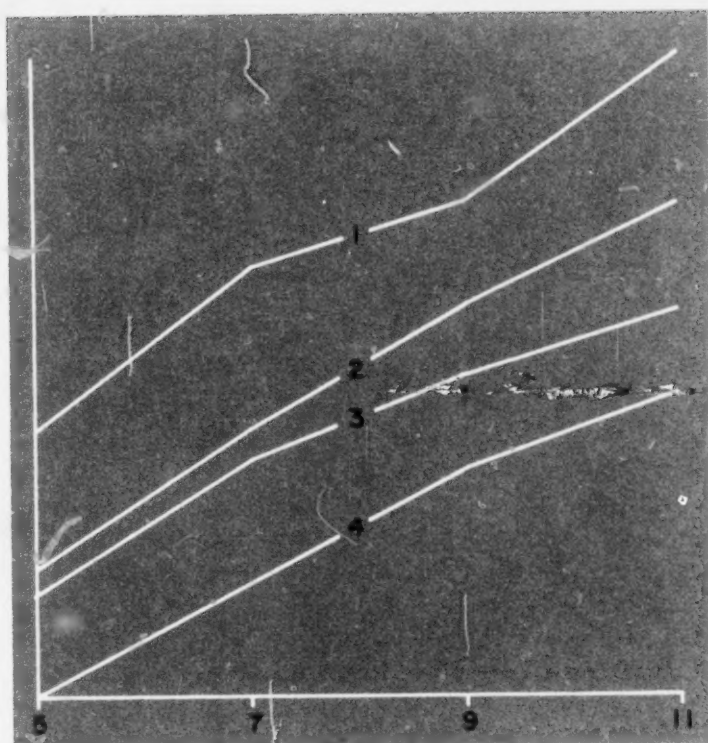
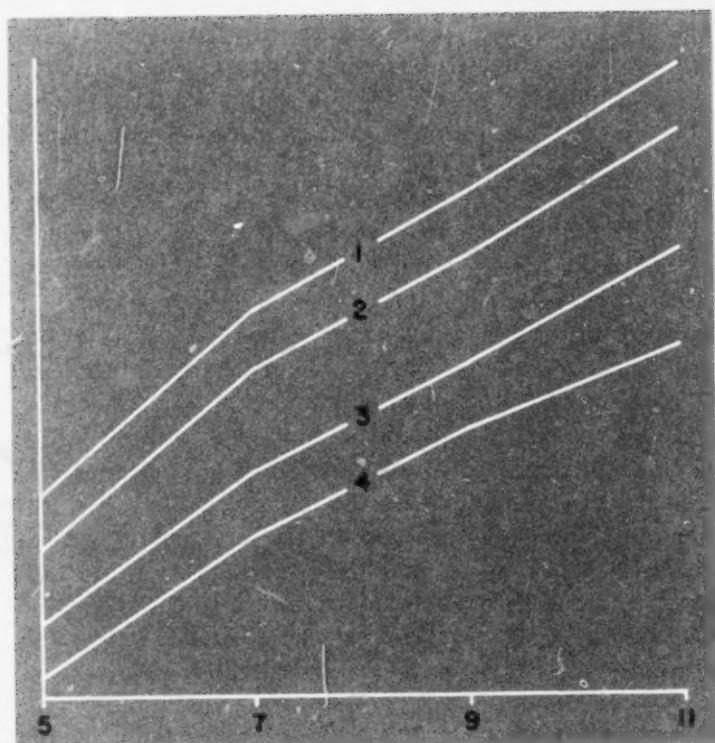
followed up through the first year after high school graduation, in the spring of 1969. The Academic Interest Measures (AIM) were also administered to the eleventh graders, along with an extensive array of questionnaires. Norms for AIM were obtained, and studies of its reliability and construct validity were conducted. These studies suggest that use of AIM as a criterion instrument may be warranted. Studies of its predictive validity are being completed, and analysis of these data should indicate whether AIM is a useful instrument for guidance. The research should be of interest to guidance counselors, curriculum evaluators, and test development specialists. CEEB. RB-67-39. *Educational and Psychological Measurement* (27, 1085-1089, 1967).

*Follow-Up Study of a National Sample of High School Seniors, Phase Three.*

BARBARA PITCHER, JAMES F. WOHLHUETER

Continuing research is needed to keep the Scholastic Aptitude Test (SAT) and the Preliminary Scholastic Aptitude Test (PSAT) up-to-date. The purpose of this study was to obtain information that could be useful in refining norms for the tests. In general, it sought to find out how the high school PSAT scores of students in the 1960s have related to their college performances. Questionnaires were used to sample 2,423 high school students who had taken the PSAT in 1960. One of many findings was that at least 62 percent of the college entrants (58 percent of the boys and 66 percent of the girls) who scored in the top five percent of PSAT scores for all 1960 high school seniors went on to graduate from college within four years. The corresponding percentage for the top 30 percent who graduate was 43, and for those scoring below the 70th percentile the percent graduated was 25. The study follows an earlier study of part of the same group that related their PSAT scores to their records as college freshmen. The report should be of interest to administrators and planners in secondary and higher education, high school counselors, college admissions and guidance personnel, researchers in secondary and higher education, and test development specialists. SR-68-88.





Mean Scale Scores on STEP Reading in Grades 5, 7, 9, and 11

1 Girls Academic  
2 Boys Academic

3 Girls Nonacademic  
4 Boys Nonacademic

1 White Academic  
2 Negro Academic

3 White Nonacademic  
4 Negro Nonacademic

The ETS Growth Study has followed the educational progress of more than 45,000 students for varying time periods over eight years. These charts illustrate some findings. The curves show, for example, an overall general rise in scholastic achievement for boys and girls, and for whites and blacks, in academic and nonacademic curricula, from grades 5 through 11. However, the curves also reflect group differences, with the second chart showing the familiar differences between whites and blacks in achievement. Findings of this type have been used to support contentions that blacks are genetically deficient in intellect. The first chart, however, shows the same ordering of differences in achievement between boys and girls. As boys are not generally regarded as inherently inferior to girls intellectually, the first chart underlines the weakness of theories of racial inferiority based on standardized test differences of the type shown in the second chart. (See Hilton, *A Study of Academic Prediction and Growth in Grades 5 through 12*.)

## Higher Education

### *The Scope of Organized Student Protest in 1967-1968.*

RICHARD E. PETERSON

By the end of the decade college student unrest had escalated to the point where many officials responsible for higher education in America considered it to be their major problem. This is the second of two studies by Dr. Peterson on student protest phenomena. It assays student protest during 1967-68 and then compares these results with his earlier study of 1964-65. Questionnaires were sent to officials of 1,000 accredited four-year colleges in the country. A few findings: During 1967-68, issues pertaining to instruction, faculty, and freedom of expression rarely evoked student activism. Issues bearing on personal freedoms and student participation in the governance of the college somewhat more often generated protest. The Vietnam war was the single issue most frequently cited by the deans as having triggered student activism. Protest varied according to type of institution—draft protests took place at 55 percent of the independent universities in contrast to 20 percent or fewer of the Catholic, teacher-training, and technical institutions. The largest proportions of student bodies were protesting issues involving administrative controls over their personal lives, while proportions protesting the draft and war-related recruiters were quite small. A comparison of the two studies, 1964-65 versus 1967-68, was particularly interesting. Campuses experiencing Vietnam war protesting almost doubled during this interval. The demand for a larger student role in campus governance increased substantially. Civil rights activism on campuses with predominantly white student enrollments declined significantly. Black student demands for black history and other studies relevant to their new self-conception increased. The proportions of activists within student bodies did not increase, though the number of colleges reporting student Left groups almost doubled, from 26 percent in 1965 to 46 percent in 1968. The study

concludes that while ethical responsibility remains a crucial issue, the potential of the student movement for renewing American life "fairly staggers the imagination." The study is available from the Institutional Research Program for Higher Education. An abridged version appeared in Foster and Long (Eds.). *Protest! Student Activism in America*. Morrow, 1969.

### *Conversations Toward a Definition of Institutional Vitality.*

RICHARD E. PETERSON, DAVID E. LOYE

Many pressures today are placing colleges and universities in a new position of social accountability. If a college is to serve its society, it would seem it must in some sense be "vital." But how is one to define "institutional vitality"? And once some conception of institutional vitality is developed, can a satisfactory instrument be constructed to assess colleges and universities against such a criterion? These questions were explored during the early research on the new Institutional Functioning Inventory: To tap a wide range of opinions, a questionnaire was sent to 1,305 people in higher education across the country. Then to explore the questions in greater depth two panels of experts were gathered at ETS for extensive discussions, which were tape recorded and edited for analysis. Both survey results and the edited discussions have been published in *Conversations Toward a Definition of Institutional Vitality*, available from the Institutional Research Program for Higher Education at ETS. In addition to ETS staff experts, discussants included Morris Keeton, Vice President of Antioch College, Warren Bryan Martin of the Center for Research and Development in Higher Education in Berkeley, Otis Singletary, Vice President, American Council on Education, Earl McGrath, JB Lon Hefferlin and Hans Flexner of the Columbia University Institute of Higher Education, and sociologists Harold and Pauline Pepinsky of Ohio State University. The research should be of interest to trustees, administra-

tors, researchers, and foundations officials concerned with higher education. Kettering Foundation.

### *Institutional Vitality Study.*

RICHARD E. PETERSON, RODNEY T. HARTNETT, JOHN A. CENTRA

What makes some colleges and universities grow and contribute in major ways to the individual and to society? Why do others merely hold their own or even go downhill? What, too, might account for peaks as well as slumps in significance for a college over the long stretch of time? To answer such questions, and to develop a practical means of measuring whatever may constitute "vitality" in an institution, this project was launched in 1967 in cooperation with the Institute of Higher Education of Teachers College, Columbia University, with funding from the Kettering Foundation. An extensive survey was made to find out what "institutional vitality" means to people who are concerned with the totality of a college or university—presidents, deans, student editors, writers. Conferences with experts, to probe the problems of what should be measured and how, were also held. Out of this experience a new instrument called The Institutional Functioning Inventory was developed. It has been extensively pretested and is now available with a preliminary technical manual from the Institutional Research Program for Higher Education at ETS.

### *Trustee Study.*

RODNEY T. HARTNETT

The traditional role of the trustee as the one who is ultimately responsible for charting the course of his college is being widely challenged. As crisis after crisis erupts on the American college campus, the trustee is often caught in the middle with both faculty and students demanding a greater share in the governing of colleges and universities, and being a more remote figure he is usually seen in terms of negative stereotypes. This study



sought to go beneath the stereotype and find out who the trustees really are, what they do in their roles as trustees, and how they feel about current issues and problems in American higher education. A questionnaire was mailed to a national sample of over 9,000 trustees. One of many findings was that trustees are generally somewhat cautious regarding the notion of academic freedom—generally they prefer to exclude even members of the faculty from decisions having to do with the academic program of the institution. In these and other areas, the attitudes of trustees differed markedly from those of the faculty. However, the study also showed in considerable detail that it is naive to speak of the "college trustee" as being any one type of person. It found a great deal of diversity between and among trustees serving on the boards of different types of institutions. The research should be of general interest to educators, particularly college administrators and faculty members. It has been published by ETS under the title *College and University Trustees: Their Backgrounds, Roles, and Educational Attitudes* (1969).

#### *The Trustee in North Carolina.*

JUNIUS A. DAVIS

In the spring of 1968 some 1,600 trustees of colleges and universities in North Carolina were sent a questionnaire. The questions dealt with who the trustees are, the nature of their service to their college or university, and where they stood on a number of crucial, current issues facing higher education. From the responses of 800 a number of interesting findings emerged. For example, 88 percent of the responding trustees were male. Only 7 percent were below the age of 40. Practically all trustees for traditionally white institutions were white, while for traditionally black institutions only half were black. Trustees were overwhelmingly Protestant, only 3 percent being Catholics or Jews. Their median annual income was \$25,000. In public institutions about one-half and in private institutions about one-third were business executives. About 90 percent of the trustees of the public institutions reported they were Demo-

crats; however, their preferences among Presidential candidates at the time of the survey were Richard Nixon first, Nelson Rockefeller second, then Eugene McCarthy, Lyndon Johnson, and Ronald Reagan. The majority—80 percent—were serving on a college or university trustee board for the first time. Only 7 percent felt faculty members should be involved in the selection of a college president. Two-thirds felt loyalty oaths for professors were reasonable. More than three-fourths favored official screening of campus speakers and almost half favored administrator control of the content of student newspapers. The research should be of interest to psychologists and sociologists researching higher education, and to others concerned with college governance.

#### *Agreement with Views of Martin Luther King, Jr., Before and After His Assassination.*

RODNEY T. HARTNETT,  
CAROL U. LIBBY

Prior to the assassination of Dr. Martin Luther King, the trustees of a number of American colleges and universities had received a lengthy questionnaire covering attitudes on many issues in connection with the College Trustee Study. The fact that a sizable number of the questionnaires were returned before and a sizable number after Dr. King's death suggested the possibility that this socially traumatic event might be reflected in a change in questionnaire responses. Comparison of the questionnaire data showed that there had, indeed, been a pronounced apparent change in the political and social views of this group. Whereas before Dr. King's death, many trustees expressed attitudes at variance with King's, questionnaires received after his death showed a liberalized attitudinal change in the direction of Dr. King's views. Significantly, this change in attitudes occurred more often among Republicans than among Democrats, and more often among conservatives than liberals—Democrats and liberals, it was assumed, being closer to Dr. King's views originally. The research should be of interest to psychologists, sociologists, political scientists, and other

social and behavioral scientists. The findings are in keeping with certain events and group behavioral patterns following President Kennedy's assassination. RB-68-35.

#### *Academic Growth in Predominantly Negro and Predominantly White Colleges.*

JOHN A. CENTRA, ROBERT L. LINN,  
MARY ELLEN PARRY

Some critics feel the value of an education at most Negro colleges is slight. In this study seven predominantly white colleges were matched with seven predominantly Negro colleges according to purposes, type of control, and ability level of students. Then the two groups were compared in terms of the scores of students who had taken the Area Tests of the Graduate Record Examinations Institutional Testing Program as college seniors and the Scholastic Aptitude Test as entering freshmen. The purpose was to see whether under these circumstances one group of colleges would show more academic improvement of its students than the other. No significant differences were found between the two groups of colleges in GRE Area Test scores in Social Science or Natural Science when using SAT scores as covariates. Moreover, inspection of regression lines for the data from the colleges also supported the conclusion that neither group of colleges was any more or less effective in encouraging academic growth. The research should be of interest to faculty and administrators in higher education, agencies interested in Negro education, and researchers in higher education. A report appeared in the January, 1970 issue of the *American Educational Research Journal*. RB-69-39.

#### *Differences in Selected Attitudes and College Orientations Between Black Students Attending Traditionally Negro and Traditionally White Institutions.*

RODNEY T. HARTNETT

Over the past few years many predominantly white colleges and universities have been going out of their way to search for qualified Negro students. However, this practice is viewed with mixed feelings by some observers

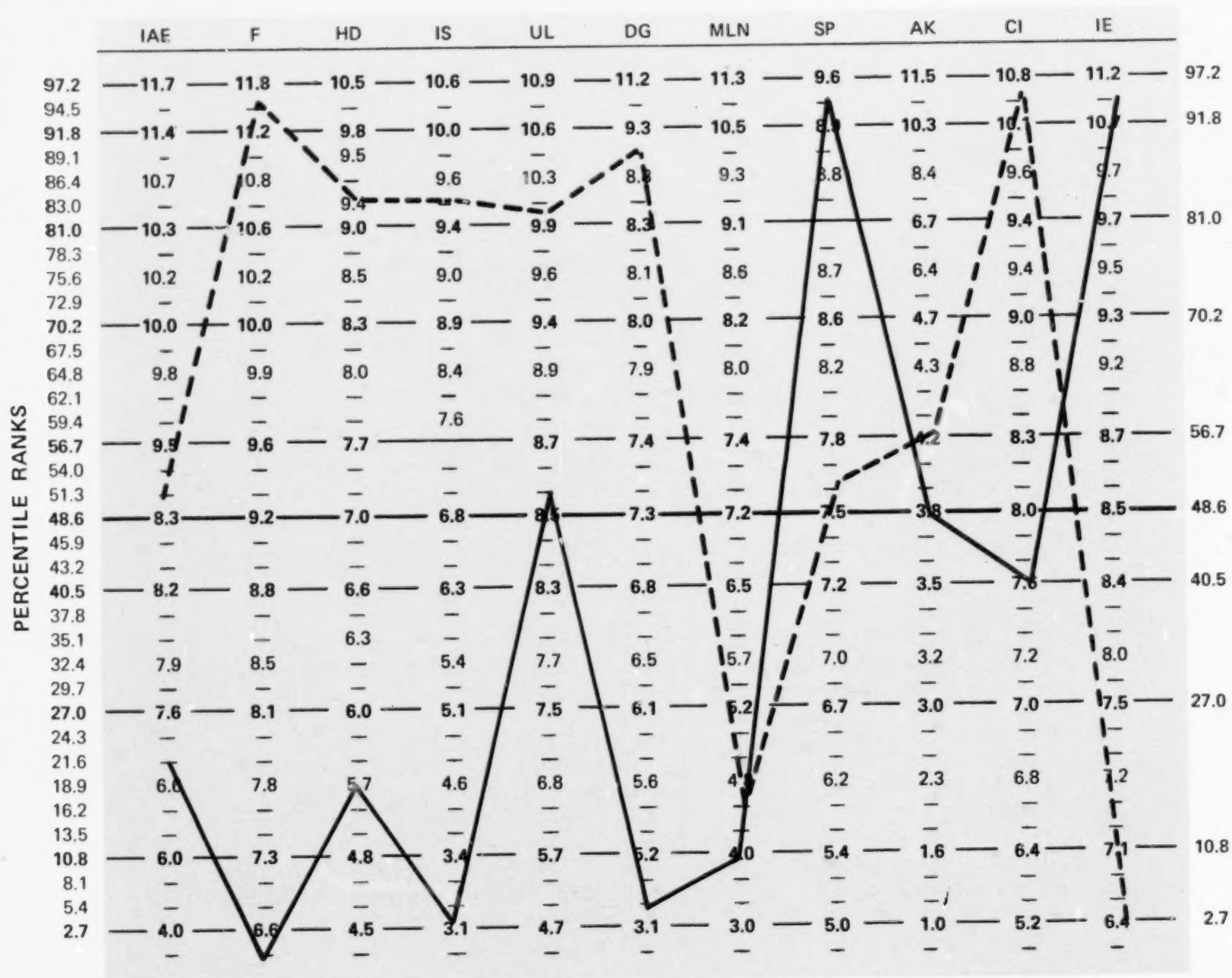


#### IFI SCALES

1. Intellectual-Aesthetic Extracurriculum (IAE): the extent to which activities and opportunities for intellectual and aesthetic stimulation are available outside the classroom.
2. Freedom (F): the extent of academic freedom for faculty and students as well as freedom in their personal lives for all individuals in the campus community.
3. Human Diversity (HD): the degree to which the faculty and student body are heterogeneous in their backgrounds and present attitudes.
4. Concern for Improvement of Society (IS): the desire among people at the institution to apply their knowledge and skills in solving social problems and prompting social change in America.
5. Concern for Undergraduate Learning (UL): the degree to which the college--in its structure, function, and professional commitment of faculty--emphasizes undergraduate teaching and learning.
6. Democratic Governance (DG): the extent to which individuals in the campus community who are directly affected by a decision have the opportunity to participate in making the decision.
7. Meeting Local Needs (MLN): institutional emphasis on providing educational and cultural opportunities for all adults in the surrounding communities.
8. Self-Study and Planning (SP): the importance college leaders attach to continuous long-range planning for the total institution, and to institutional research needed in formulating and revising plans.
9. Concern for Advancing Knowledge (AK): the degree to which the institution--in its structure, function, and professional commitment of faculty--emphasizes research and scholarship aimed at extending the scope of human knowledge.
10. Concern for Innovation (CI): the strength of institutional commitment to experimentation with new ideas for educational practice.
11. Institutional Esprit (IE): the level of morale and sense of shared purposes among faculty and administrators.

A relatively new instrument, The Institutional Functioning Inventory (IFI), has been developed by the ETS Higher Education Group to aid colleges and universities in the difficult task of self-assessment. A standard questionnaire is given to faculty, administration, in part to students, or to any other group whose rating of the institution is desired. Extensive research established that these 11 dimensions are the most relevant for rating purposes. (See Peterson, *Institutional Vitality Study*.)

# INSTITUTIONAL MEANS



The possibilities of the Institutional Functioning Inventory (IFI) are perhaps most quickly evidenced by the clear and carefully quantified profiles the instrument makes possible. Means for responses to questions relating to each of the 11 institutional dimensions are plotted on a standard grid. Shown here are the contrasting profiles of two colleges based on the perceptions of each institution by its faculty. The profiles are of an armed-services academy (solid line) and a selective liberal arts college (dotted line). (See Peterson, *Institutional Vitality Study*.)

who are concerned that siphoning off talented black students will hurt the quality of the Negro schools. The purpose of this study was to see what educationally relevant differences exist, if any, between two groups of black collegians—those who enter traditionally Negro colleges and those who enter integrated colleges. College Student Questionnaires, Part I data were obtained from 3,104 students at 9 traditionally Negro colleges and 323 students at 21 integrated colleges. In addition, data on scores from the verbal sections of the Scholastic Aptitude Test (SAT) were obtained for a sample of each of these groups. Perhaps the most notable single finding was that black students entering integrated schools were found to have higher SAT-V scores, to be more independent, more liberal, more concerned with social injustice, and to aspire to more years of formal education. Many attitudinal differences between the two groups were also found to be highly correlated with SAT-V scores. The research should be of interest to researchers in higher education, and to those concerned with the aspirations of the black collegian. RB-69-48.

*Development of the Questionnaire on Student and College Characteristics.*

JOHN A. CENTRA

Both colleges and students seeking to enter them have a problem in common. Underneath the mountain of words and pictures directed to presenting the college in catalogues and little booklets, both colleges and students seek to grasp what the college is *really* like. This study developed the Questionnaire on Student and College Characteristics (QSCC) for simplifying the process. Using upperclassmen as respondents, the questionnaire gathers three types of information about a specific college

—student perceptions of the college and its learning environment, student behavior patterns, and student characteristics, such as religious and ethnic mix, or interest patterns. The instrument was pretested at eight colleges, revised, and then administered at over 200 institutions in the fall of 1968. Information from the QSCC may be used by colleges to describe themselves in *The College Handbook*. Currently research is under way to identify dimensions that will describe meaningful differences among colleges; eventually the results will be related to college environment studies. The finished instrument and the research related to differences among colleges should be of interest to high school guidance counselors, college admissions officers, prospective college students, and researchers in higher education. CEEB. RM-68-11.

*College Freshman Attitudes Toward Cheating.*

JOHN A. CENTRA

This study sought answers to two questions: (1) what are the characteristics of students with lenient attitudes toward cheating, and (2) do attitudes toward cheating vary according to different types of schools? The student sample consisted of 1,500 entering freshmen at 37 colleges, the institutional sample of 119 colleges and universities. It was found that students with lenient attitudes toward academic cheating shared similar attitudes about cheating in government and industry. The cheating-lenient student also tended to be less academically motivated, to have fewer artistic-literary interests, and to come from a lower socioeconomic background than the cheating-strict student. As for school differences, it was found that students at institutions that were fairly selective, had all-female enrollments, and were small in

size had the strongest attitudes against cheating. The study suggests that the kind of student who enrolls in an institution is an important determinant of the peer climate, which, in turn, sets the school's characteristic attitudinal climate toward cheating. The study, which was reported in the January, 1970, issue of the *Personal and Guidance Journal*, should be of interest to counselors in colleges and to researchers in higher education. RB-69-24.

*Attitudes and Secondary School Backgrounds of Catholics Entering College.*

RODNEY T. HARTNETT, JOHN A. CENTRA

Recent research has indicated that a cumulative Catholic education—that is, consistently attending Catholic schools over a number of years—produces adults who are not only more religious than customary but whose attitudes also reflect more social concern. This study used the College Student Questionnaires, Part I, to obtain data for comparing the backgrounds and attitudes of 1,700 men enrolled at six Catholic and seven non-Catholic colleges. The most notable finding was that students who chose the Catholic colleges reflected more social concern than those who chose the non-Catholic colleges. This suggests that the finding of more social concern is not necessarily the result of the cumulative effect of attending Catholic primary and secondary schools and colleges, but could be the result of a choice influenced by the nature of Catholic colleges. The research should be of interest to Catholic college administrators and researchers in higher education. It questions some of the conclusions reached by Greely and Rossi in *The Education of American Catholics*. A report appeared in *Sociology of Education*, Spring, 1969. RB-68-24.



## Graduate and Professional Education

### *Investigation of Graduate Admission Policies and Procedures.*

RICHARD L. BURNS

Like all other aspects of education, graduate schools are undergoing considerable pressure for change. The purpose of this project is to gather information on graduate admissions policies, procedures, and trends in every graduate school in the country, and then to make the results widely known. The information will be used (1) to aid the Graduate Record Examinations Board and ETS in planning programs and other activities in graduate admissions; (2) to make available to all graduate schools up-to-date information on procedures and policies being followed by other schools throughout the country; and (3) to stimulate thinking in the graduate schools about admissions procedures and policy improvements. During Phase I a survey will be made of all graduate schools. Phase II will involve visits to a limited number of graduate schools to develop case studies of admissions policies and procedures. When completed, the research should be of interest to administrators and graduate school admissions personnel. GRE Board.

### *GRE Creativity Project.*

JONATHAN R. WARREN

In various ways graduate schools seek to identify creative students who may go beyond the traditional learning in any field to enlarge or improve it. It is, however, much easier to understand creativity informally than it is to measure it formally with standardized instruments. The purpose of this project is to explore the measurement of creativity at the graduate level by first defining creativity as a matter of productivity, originality, and independent thinking, and then attempting to identify it in students by interrelating various measures to the Graduate Record Examinations. Measures being evaluated include a biographical inventory, and a revised Independent Activ-

ities Questionnaire, which is a relatively recent standardized measure of some aspects of creativity developed by ETS. Criteria include peer ratings of the degree to which a fellow graduate student seems creative, graduate school grades, and situational tasks that require some degree of creativity in their resolution. Data have been collected on students in electrical engineering, psychology, and English at five graduate schools in each field. When completed, the research should be of interest to graduate admissions personnel, test developers, and psychologists working in the area of creativity. GRE Board.

### *Alternative Methods of GRE Advanced Tests.*

J. GARY LUTZ, SUSAN F. FORD

An important problem in test development is the equating of tests across forms. That is, most tests must be changed from administration to administration. But each new form, though changed and containing fresh items, must still measure the same abilities and also maintain the same overall proportion of level of difficulty. Equating, therefore, becomes quite a complex problem. The purpose of this study is to see whether what is known as common-item equating of the Graduate Record Examinations (GRE) Advanced Tests is superior to a method using the verbal and quantitative sections of the GRE Aptitude Test. The study is designed to answer three questions: one, does equating through aptitude tests, in general, yield significantly different results from those yielded by common-item equating? Two, does stability of equating through the aptitude tests compare to that of common-item equating across various administration months? And three, how does the stability of equating through aptitude tests compare to common-item equating across various educational levels of candidates? When completed, the research should be of interest to test developers. GRE Board.

### *GRE Cooperative Study of Predicting Graduate School Success.*

GERALD V. LANNHOLM, WILLIAM B. SCHRADER, GARY L. MARCO

This study explored possible methods for increasing the predictive validity of the Graduate Record Examinations (GRE) Aptitude and Advanced Tests. Undergraduate grades for a number of doctoral candidates were obtained and also ratings by professors of their current excellence, achievement, and promise in graduate school. The six academic fields studied were chemistry, physics, English, philosophy, history, and psychology. It was found that undergraduate grade-point average (GPA), which was available for 12 groups, was a reasonably good predictor in four of these groups, but a poor predictor in five groups. When undergraduate GPA was combined with GRE test scores using *judgmentally determined* weights, the weighted total yielded better prediction than undergraduate GPA alone in all 12 groups, (with a gain of .14 or more in six of the groups). The results offer some support for the use of a priori weighting—or placing more value on some test results than on others—in combining predictors. The research should be of interest to graduate admissions personnel. GRE SR-68-3.

### *The Use of GRE Scores and Other Factors in Graduate School Admissions.*

GERALD V. LANNHOLM

This study was undertaken to gain information about the use of Graduate Record Examinations (GRE) scores and other factors in admitting practices for graduate schools. Departments in 30 universities were sent questionnaires. The results of the survey indicated that all respondents collected more than one type of information about applicants. Approximately 40 percent had established formal admissions standards in terms of minimum GRE scores required. However, only a relatively small percent (6.6) used test scores as the principal factor in admissions decisions, with a range of mini-

mum scores from 450 to 750 for verbal ability, and 450 to 700 for quantitative ability. Most gave the highest ranking to the undergraduate record. While total agreement on the value of letters of recommendation wasn't found, 33 percent indicated that they would never admit applicants with high GRE scores but poor letters. However, 74 percent reported that they would sometimes admit applicants with low GRE scores but good letters. GRE SR-68-4, October, 1968.

*Validity of GRE Advanced Tests for Predicting Law School Success.*

WILLIAM B. SCHRADER,  
BARBARA PITCHER

The purpose of this study was to determine whether the Advanced Tests of the Graduate Record Examinations (GRE) will improve prediction of law school grades when used along with the Law School Admission Test (LSAT) and undergraduate grade averages. It is plausible, for example, that an undergraduate who performs well on tests in his major field will be a better law school student than one whose GRE Advanced Test performance is mediocre or poor. Advanced Tests were given to about 900 students in six law schools in the fall of 1966. Their first-year law school grade averages were also obtained and validity coefficients computed for each law school. Preliminary findings indicate that GRE scores do not add appreciably to law school grade prediction over use of the LSAT and undergraduate average or grades alone. However, GRE scores do correlate moderately with law school grades in the relatively homogeneous groups that were studied. LSAT Council.

*Continuation of Criterion Study: LSAT*

ROBERT L. LINN

This study sought to find out to what extent grades assigned by law professors to students' essays are dependent upon the content of a student's answers, the way he approaches the problem, and certain extraneous factors, such as the number of words or the number of polysyllabic words he uses. Although initial results indicated that essay grades were quite predictable from

simple essay length, a more refined analysis indicates that it is only the number of words that are expended on issues judged to be of a major importance that accounts for this predictability. Sponsored by the Law School Admissions Test Board, this research should be of interest to law schools and to any other group interested in essay tests. It provides a better understanding of the factors influencing the assignment of grades in essay examinations and relates to a large body of research on the reliability and validity of essay tests. LSAT Board.

*Interpreting Performance of Foreign Law Students on LSAT and the Test of English as a Foreign Language. (TOEFL)*

WILLIAM B. SCHRADER

The purpose of this study was to investigate two questions: How effective are LSAT and TOEFL scores in predicting language effectiveness and academic performance of foreign law students? And to what extent do foreign students show marked improvement in performance on the tests as the result of a relatively brief orientation program which includes special instruction in English for some students? The LSAT and TOEFL were given at the beginning and end of an eight-week intensive orientation program during the summer of 1968. Additional data are being gathered for analysis. LSAT Council, Orientation Program in American Law, and CEEB.

*Factor Analysis of the LSAT.*

ALFRED B. CARLSON

The purpose of this study was to gain a better understanding of the abilities measured by the Law School Admission Test (LSAT). Specifically, the questions were: What abilities are current LSAT items measuring? And are the abilities measured by the afternoon tests (Writing Ability and General Background) different from those measured by the morning test? Seven factors were identified in the morning LSAT: reading comprehension, figure classification, tabular data interpretation, graphical data interpretation, numerical recall, verbal inductive reasoning, and a seventh uninterpreted

factor. Conclusions drawn on the basis of the findings of the study should be of interest to test developers. LSAT Council.

*Summary of LSAT Research.*

ROBERTA B. FLESHER

Over the past 20 years, numerous research studies of various questions relevant to the development or improvement of the Law School Admission Test (LSAT) have been carried out. This project summarizes 27 of these studies, and should be of interest to law school admissions officers, new LSAT Committee members, educators, and psychologists. LSAT Council.

*Special Item Analysis of the Admission Test for Graduate Study in Business for Candidates Sponsored by the Consortium for Graduate Study in Business for Negroes.*

WILLIAM R. COWELL

The Consortium for Graduate Study in Business for Negroes was formed in 1966 to encourage Negro undergraduates aiming toward a business career to enter graduate schools of business. The purpose of this study was to gain a better understanding of the general nature of candidates being selected by the Consortium in comparison with non-Consortium candidates. Admission Test for Graduate Study in Business responses for the two groups were compared on an item-by-item as well as total-score basis. It was found that the two groups differed significantly on verbal and quantitative scores, the Consortium group scoring higher than the non-Consortium group on verbal. Significant differences were found on 12 of 90 verbal items and 14 of 75 quantitative items, there being generally a remarkable similarity in the distribution of wrong responses among the options. Among quantitative items, those involving percentages were found to be significantly more difficult for the Consortium group. The test was found to be about equally reliable for both groups. The research should be of interest to graduate business school admissions personnel, researchers in higher education, and educators concerned with social action. ATGSB.



### *ATGSB Grade Grouping Study.*

ROBERT F. BOLDT

Among those concerned with the admissions problems of graduate schools of business, there is widespread dissatisfaction with grade-point average (GPA) as an indicator of academic performance. One aspect of this dissatisfaction is that the grade-point average is a conglomerate, which, by averaging grades, levels off strengths on one hand and on the other hand masks weaknesses. The purpose of this study is to determine, on a pilot basis, whether there is some reductive alternative other than the single GPA score. For example, can a few general abilities of potential importance for comparison purposes be discerned underlying groups of courses? If such a relationship of underlying abilities to groupings of courses could be discerned, it might provide a more adequate means of interpreting grades and test scores. This study is exploring such questions factor-analytically with data from the graduating classes of two graduate schools of business. To date, analysis of data from the schools yields a single factor that correlates highly with grades in mathematics and quantitative studies generally and with the quantitative portion of the Admission Test for Graduate Study in Business. If this finding proves general, the opinion that grades are a conglomerate may be questioned. Hence, the study should be of interest to testing theorists and psychometricians, as well as to graduate business school admissions personnel. ATGSB.

### *Discrepant Predictors Study for the Admission Test for Graduate Study in Business.*

ROBERT F. BOLDT

One familiar dilemma that admissions officers face is how to decide among candidates whose test scores and grade-point averages do not clearly point to the same conclusion as to the candidate's potential ability. A similar problem is faced by the researchers who, in carrying out validity studies, must develop a predictive system that works no matter what configurations of

test scores and undergraduate grades are available. The purpose of this study was to see whether the methods of the validity study might be used by the admissions officer. In general, the indications are that the regression—equation approach used in validity studies does work reasonably well and can be used to resolve the conflicting information with which the admissions officer deals. The research should be of interest to graduate school admissions officers and test developers. ATGSB Brief No. 2, June 1969.

### *Criterion Study: ATGSB.*

LOIS A. CROOKS

The broadening of criteria for selection purposes is a long-standing and difficult problem in testing. For example, college grades have long served as useful criteria for determining the validity of cognitive measures, such as the aptitude and achievement measures now in general use. The fact that success in school—and in life—also depends on qualities *not* measured by aptitude or achievement tests has led to persistent efforts to create useful measures of noncognitive qualities, such as those of personality or motivation. However, it has been extremely difficult to establish criteria as obvious and quantifiable as grades for validating measures of these other qualities. The purpose of this study is to identify criteria that might be useful in validating tests of noncognitive qualities that graduate schools of business most desire in their students. During the first phase of the study some desirable qualities were identified and rating scales were developed. In phase two, a biographical data form was developed and pilot studies were carried out in graduate schools of business. In phase three, noncognitive data from the biographical data form are being used along with undergraduate grade-point averages and scores on the Admission Test for Graduate Study in Business to see whether adding the noncognitive data improves the predictability of performance by graduate business school students. The research should be of interest to researchers in test development and higher education. ATGSB.

### *ATGSB New Item Type Research.*

ALFRED B. CARLSON

This study investigates the validity of three new item types for predicting success in graduate study in business. One item type, the *Common Word* type, requires the candidate to give associations for three words and then find the one word that is common to all three. A second item type, *Quantitative Comparison*, requires him to state which of two quantities is greater, to assert their equality, or to assert lack of sufficient data to reach a judgment. A third type, *Practical Business Judgment*, requires the reading of a short background piece on a business executive with a hypothetical problem, and then to make the classifications and judgments necessary to solve the problem. Item writing has been completed, test copy and a supervisor's manual are being prepared, and six schools with a total of about 1,500 students will participate in a field testing of the new approach. The research should be of interest to graduate business school admissions personnel and test developers. ATGSB.

### *ATGSB Moderator Variable Study.*

BARBARA PITCHER, HERMAN SMITH

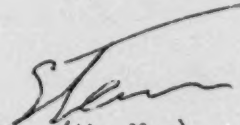
This exploratory study was undertaken to see if certain variables might act as moderators in predicting the average grades of students during their first year in graduate business school. For example, a moderator variable, *age at entrance to graduate business school*, underlies such a question as, "Are predictions based on ATGSB scores more accurate for younger students or for older students?" Other moderator variables of this nature raised such questions for research as, "Are predictions more accurate for fulltime students than for part-time students? Does predictive accuracy differ according to undergraduate major field or the excellence of candidates from particular undergraduate colleges? Would predictions using ATGSB scores and grades be more accurate if separate prediction equations were used for groups defined according to these moderator variables?" Data collected from 26 graduate schools indicated that predictions



FROM THE DESK OF G. STANLEY SALVATI

Rip:

Thought you ought to know, as one old friend to another, that the Manager of the service department of the new discount chain store is trying to hire away our electronics technicians. One of my men said this guy called him at home and offered him two bits an hour more than our base rate, plus all kinds of goodies like 25% discount, lots of overtime, and a brand new truck. Be prepared.

  
Stan (the Man)  
HiFi Service Center

*I saw this. Ron Flowers said  
he had a call.*

*Fran*

Developed by ETS in collaboration with a variety of organizations, "in-basket" tests have become a widely used measure of managerial potential. Each test is a simulation of real problems that a manager must solve in a particular job. As the name "in-basket" suggests, fictional memoranda, letters, notes, and other documents convey problems to the desk of a manager. The person taking the tests responds to these problems with various directives, which are rated for appropriateness and according to other criteria. Samples on this and the next page are from the J. C. Penney Company Service Center Manager In-Basket. (See Crooks, *In-Basket Study*.)

Manager  
Penny Co. Service  
Bradford, N. J.

Sir:

I consider it my duty to inform you that one of your trucks has been setting in the driveway next door a little too often. The woman's husband is not home. The people just moved here and she's not the friendly type or I would speak to her myself. It doesn't look good for a moral company like Penny to have their men enticed. Can you do anything to her? Her address is 114 Honeysuckle Lane. The name is Robinson.

A friend

are more accurate for younger students than for older students, and for full-time students than for part-time students. The predictive effectiveness of ATGSB scores and undergraduate grades also seems to be influenced in a small but consistent way by undergraduate major field and college. The research should be of interest to graduate school admissions personnel and test developers. ATGSB. SR-68-65.

#### *In-Basket Study.*

LOIS A. CROOKS

A sizable problem for graduate business school admissions people is deciding which students to select from among the marginal candidates—that is, applicants whose scores on the Admission Test for Graduate Study in Business, undergraduate records, and other data place them in the lowest category of possible admittees but in larger numbers than there are places available. A related problem is the selection of candidates who not only may complete the academic program successfully, but who may also have other attributes needed for success in business careers. (The problem here is pointed up by the fact that follow-up studies of business

school graduates have shown little or no relation between their grades in business schools and later career progress.) This study is investigating the possibility of using an "in-basket" test as an aid in selection for admissions. The in-basket simulates the solving of management problems, and research over the years in a variety of industrial settings indicates that in-basket performance appears to be related to actual managerial career progress. An in-basket test was administered to candidates in several graduate schools of business in the fall of 1969 together with a biographical data questionnaire and several personality measures. The data will be analyzed using grades and faculty ratings as criteria. When finished, the research should be of interest to researchers in test development and higher education. ATGSB.

#### *The Moderating Effect of TOEFL on ATGSB.*

AMIEL T. SHARON

Foreign students who want to enroll in American business schools are frequently examined with both the Test of English as a Foreign Language (TOEFL) and the Admission Test for Graduate Study in Business (ATGSB). Since no

information is available as to whether TOEFL adds to the predictive validity of the ATGSB for foreign candidates, this study combined the two tests to see how well they predict business school grade-point average. When completed, the research should be of interest to graduate business school admissions personnel. ATGSB.

#### *Selection of MBA Candidates for Later Job Effectiveness.*

THOMAS W. HARRELL

To what extent can students with management potential be identified prior to admission to a graduate school of business? This question was explored by Dr. Thomas Harrell of Stanford University's Graduate School of Business with support by the Policy Committee of the Admission Test for Graduate Study in Business (ATGSB). In this initial phase 64 men from the class of 1961 and 104 from the class of 1962 took a series of tests and questionnaires while they were in school. The ATGSB Brief No. 1, published early in 1969, describes the study and its implications. The research should be of interest to admissions personnel in graduate schools of business. ATGSB.

## *Cultural Diversity and the Disadvantaged*

### *Disadvantaged Children and Their First School Experiences: A Longitudinal Study.*

SCARVIA ANDERSON, ALBERT BEATON, JOSEPH BOYD, VIRGINIA SHIPMAN, *et al*

During the summer of 1969, about 2,000 4-year-olds in Lee County, Alabama; Trenton, New Jersey; St. Louis, Missouri; and Portland, Oregon, became part of a six-year study of educational and social programs for disadvantaged children. Study plans and operations include a close look at the families and communities of the children, in order to appraise their roles—in interaction with preschools and schools—in fostering children's mental, social, emotional, and physical

development. The ultimate goal of the study is to provide information useful for educational and social planning at federal, state, and local levels. Although the study has been described by leading social scientists as "potentially the most significant single piece of educational research in this decade," its origins were simple and straightforward: The federal government is spending hundreds of millions of dollars on Head Start and other early childhood programs, and there is a pressing need to know not only what kinds of outcomes these programs are producing, but also what can be done to improve them in the future. The present effort is an attempt to provide reliable information for program evalu-

ation. It is also designed as basic research on how human development is sequenced and structured; as social research on the pressing domestic problems of both poverty and alienation from the mainstream of society; and as a "practicum" in how to do research in the real world and how to put the knowledge gained from such research to practical use as quickly as possible. The study communities are fully involved in the research—administratively, financially, and professionally. For example, the testers, observers, and interviewers are not the usual outside researchers with clipboards, but neighborhood mothers who have been given special training. Advisers include Samuel Messick, Albert



Beaton, Walter Emmerich, and Winton Manning, ETS; Edmund Gordon, Columbia University; Marshall Smith, Trenton State College; Silvan Tomkins, Rutgers University; Melvin Tumin, Princeton University; Urie Bronfenbrenner, Cornell University; Boyd McCandless, Emory University; Alfred Yankauer, Harvard University; and Edward Zigler, Yale University. The study is supported principally by a grant from Project Head Start, Office of Child Development, HEW. PR-70-2.

#### *Classroom Variables.*

DONALD M. MEDLEY

One problem facing those working on the ETS-Head Start Longitudinal Study is how to obtain a reasonably adequate record of what happens to the children in their classrooms. What do teachers do and what do the children do that might relate to what is learned? What seems to turn them on? Or turn them off? The mission of one group of ETS psychologists, who constitute the "classroom task force," is to obtain comprehensive records of the classroom experience of the individual child and to develop ways of putting this information into a quantitative form for statistical analysis and to test various hypotheses. To do this a new instrument called the *Personal Record of School Experiences*, or PROSE, has been developed and pretested. During a pilot study local people in the various longitudinal study communities were trained to use the new instrument, acting as classroom observers. The pilot study was also used to provide data for testing computer programs that are being developed for the storage and retrieval of information gathered by PROSE. Project Head Start. RM-69-15.

#### *Planning for Innovation: A Case Study of the New York City School-Community Interaction Umbrella Program.*

GEORGE G. GORDON

A problem of great current importance throughout American society is how administrative power is to be shared between central and decentralized authorities. In education one of the most dramatic eruptions of this issue has come within New York City, where

it has proved considerably easier to recognize the need for decentralization than to effect it. This study was of a key local problem—how the newly formed, decentralized, administrative boards organized themselves to implement and oversee a variety of Title III projects in various communities. Observers attended meetings of four such decentralized boards as well as a number of interboard meetings. Copies of written communications between the boards and the Board of Education office monitoring the project were also obtained, and these inputs were analyzed. The study indicated that the boards spent most of their time defining their role, their responsibilities, and their authority. However, progress was made overall, and although the atmosphere was often abrasive and hostile, the school and community people involved were generally hopeful about the potential for effective school-community interaction. The research should be of interest to educational planners for urban areas, school administrators, and school boards.

#### *Test and Job Performance of Various Subgroups Within Specified Occupations.*

JOEL T. CAMPBELL, RONALD L. FLAUGHER, LEWIS W. PIKE, DONALD A. ROCK

This study sought to determine the relationships among aptitude tests, criteria of job performance, and background factors for Negro groups in comparison with white groups. One finding was that Negro and white supervisors evaluate Negro and white subordinates on quite different bases. Another finding was that the factor structure of aptitude tests for Negroes and whites shows some differences, but overall shows more similarities than differences. Though similar studies have been done, the data for this project were collected from a larger sample than in previous studies. Results should be of special interest in demonstrating the difficulties of adequately validating tests for diverse ethnic groups. In relation to education, they also raise some practical questions about the assignment of course grades to students from different ethnic groups. The research should be of

interest to personnel and measurement specialists, and those generally concerned with equal opportunity. Ford Foundation. PR-69-4. PR-69-5. PR-69-6.

#### *Evaluation of the Integration of Evanston District 65 Schools.*

DANIEL P. NORTON, JAYJIA HSIA

Such studies as the giant cross-sectional Educational Opportunities Survey have shown that generally speaking, across the nation, integration very importantly affects education. Many questions remain, however, of how and why this happens. Such questions can only be answered by following the same students over a period of years using the methods of the longitudinal study. This study, which started in 1967, is following the progress of children in the Evanston, Illinois, school district through June of 1971. Both cross-sectional and longitudinal analyses are being made of achievement, attitude, socioeconomic status, and other kinds of information on the children. When completed, the research will be especially relevant to studies of integration and of equality of educational opportunity because of the scarcity of studies of this nature. The completed research should be of interest to the general public, school boards, school administrators, and human growth and development specialists. Rockefeller Foundation. PR-69-10.

#### *Chicago Title I Evaluation.*

GARY L. MARCO

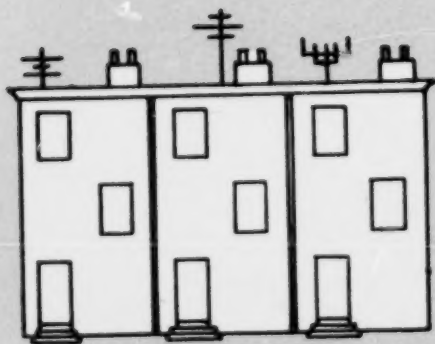
Administering Title I funds for educational innovations affecting the disadvantaged poses a number of problems. In Chicago some of these problems have been particularly acute because of the size of the school system necessary to serve the nation's third largest city. This study was designed to answer two questions of pressing importance to Chicago Title I fund administrators: How effective are the various programs funded by Title I in the city? How might these funds best be allocated to the various programs? An extensive opinion sampling of those "in the know" throughout the city and the school system was carried out. It

(continued on page 80)



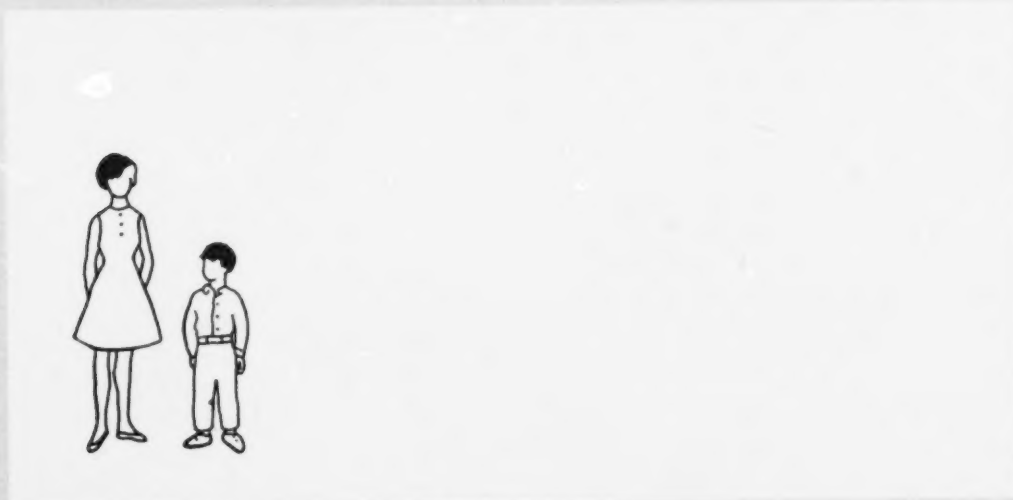
A child's self-perception and his feelings toward people and places of significance in his environment may be gauged by the *Social Schemata Task*. The child first chooses a figure to represent himself from among the four self-figures shown above: a boy, a girl, black or white. He is then equipped with a sheet of his own self-figures on pressure-sensitive labels. He pastes these self-figures onto sheets of paper that show the important people and places in his life

indicated below—mother, father, teacher, friends, both black and white, and his home and school. In this way the child may reveal relatively quickly many things of fundamental importance to his learning and well-being. Some illustrative examples of the task in use are shown overleaf. (See Anderson, *Disadvantaged Children and Their First School Experiences: A Longitudinal Study*.)

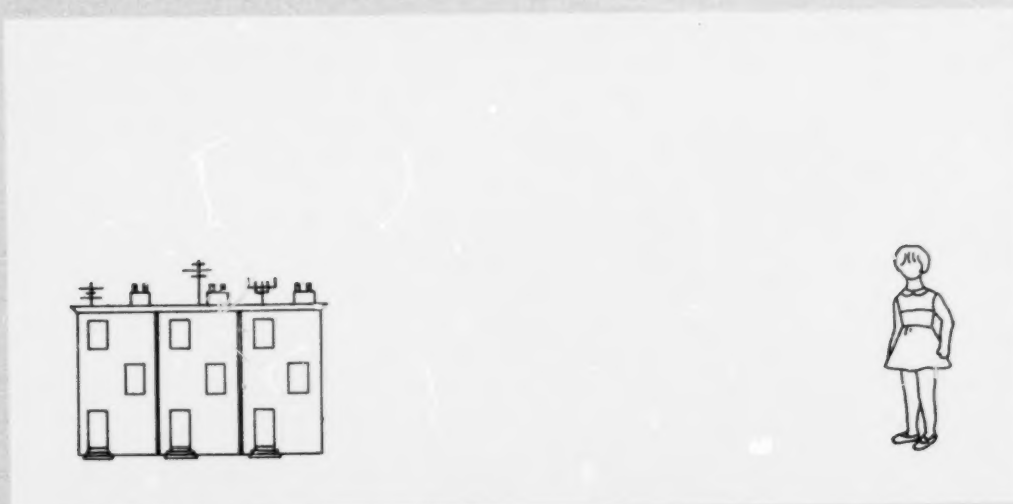


COPYRIGHT © 1969 BY EDUCATIONAL TESTING SERVICE. ALL RIGHTS RESERVED.

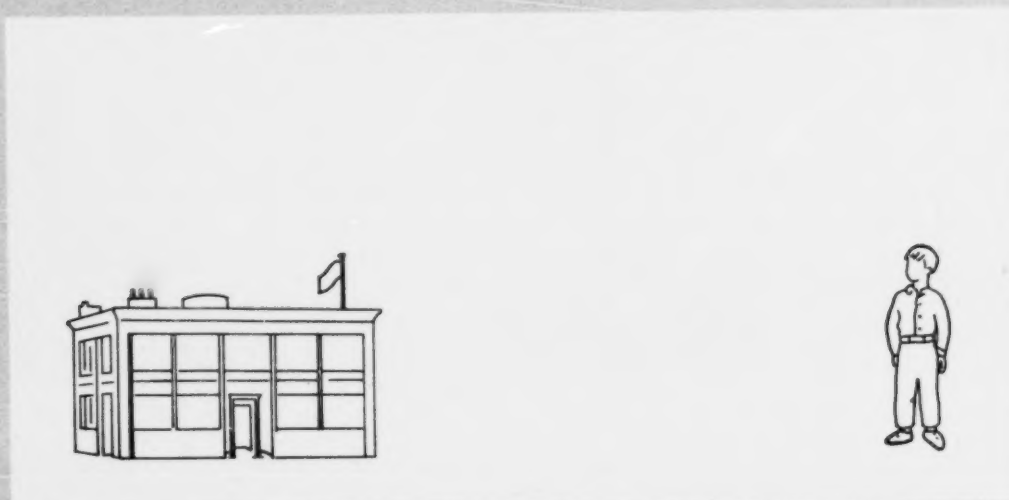
**EXAMPLE 1:** *The child places himself near the mother. This proximity of the two figures might indicate a sense of being close to his mother. An exact interpretation must, of course, take into consideration all responses to the Social Schemata Task as well as other assessments.*



**EXAMPLE 2:** *The child places herself far from home. This might indicate some degree of negative feeling toward her home.*



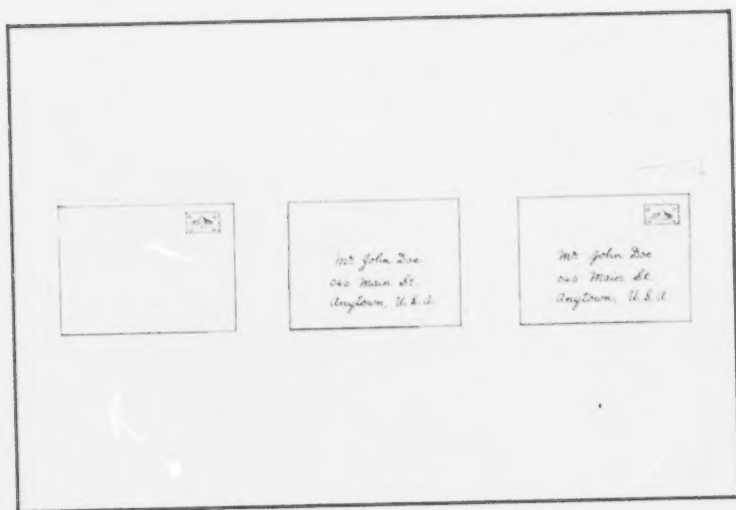
**EXAMPLE 3:** *The child places himself far from school. This might indicate reluctance to attend—or some other negative feeling regarding his school.*



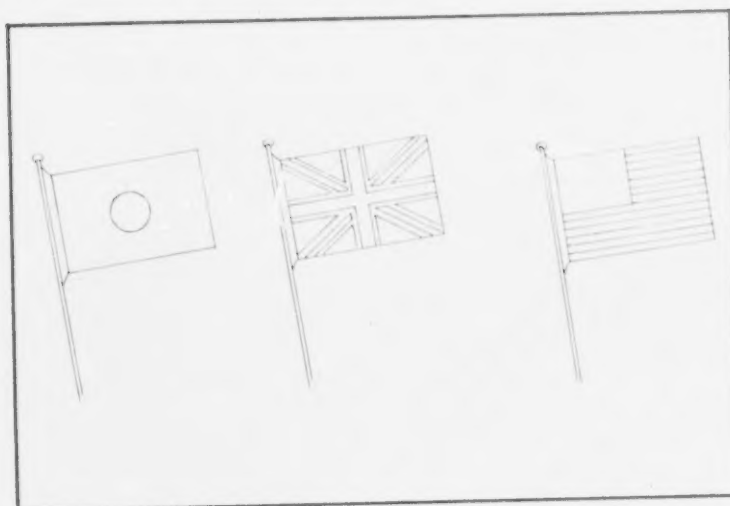


To measure a young child's domain of general knowledge the testing instrument must be sensitive to both individual and cultural differences. On this and the overleaf page are shown items from the TAMA General Knowledge Test (II). The child is shown a card with three comparison objects and asked the questions indicated. (See Anderson, *Disadvantaged Children and Their First School Experiences: A Longitudinal Study*.)

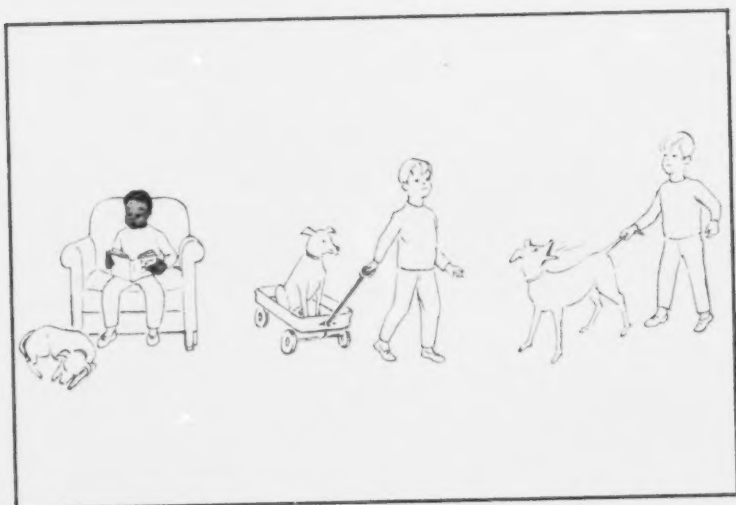
*Which letter is ready to mail?*



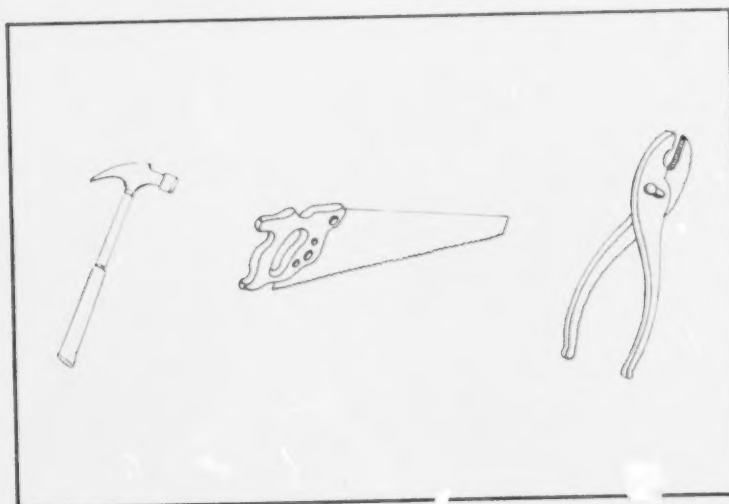
*Which is most like the flag of our country?*



*Which dog might bite?*

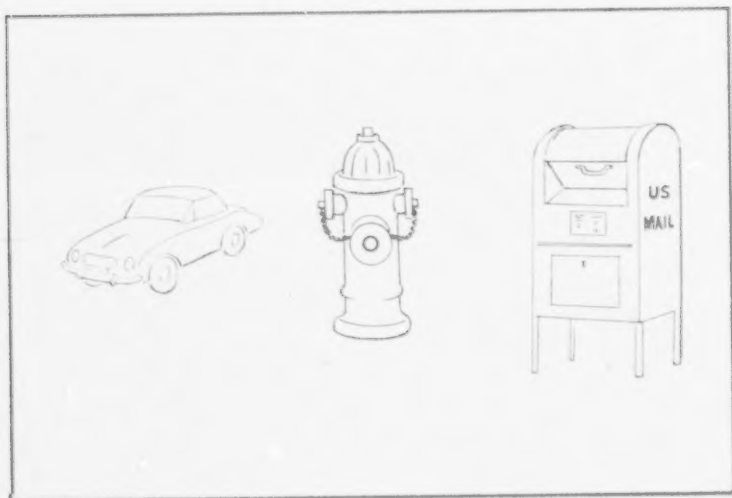


*Which could you use to cut a piece of wood?*

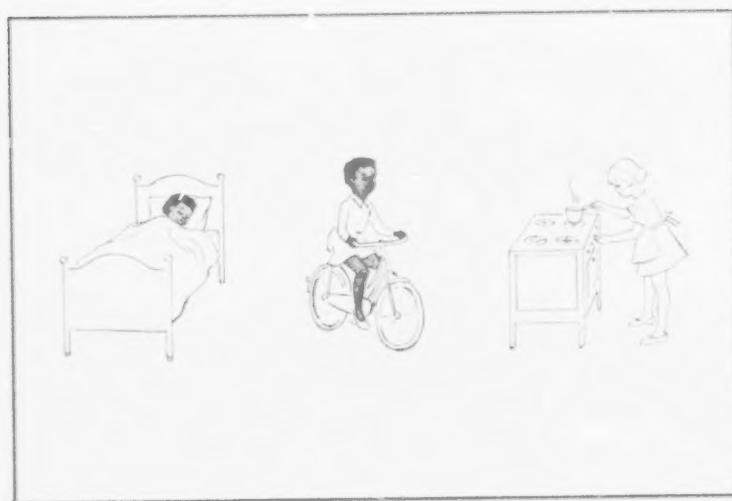


*(continued on next page)*

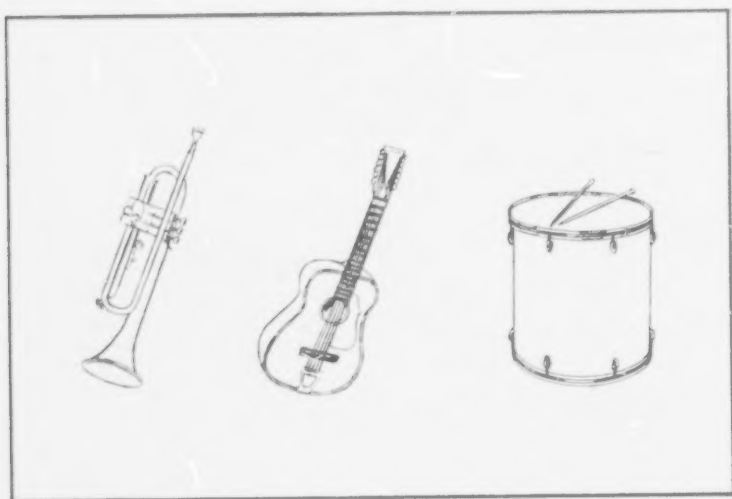
*Which of these do firemen use when they put out a fire?*



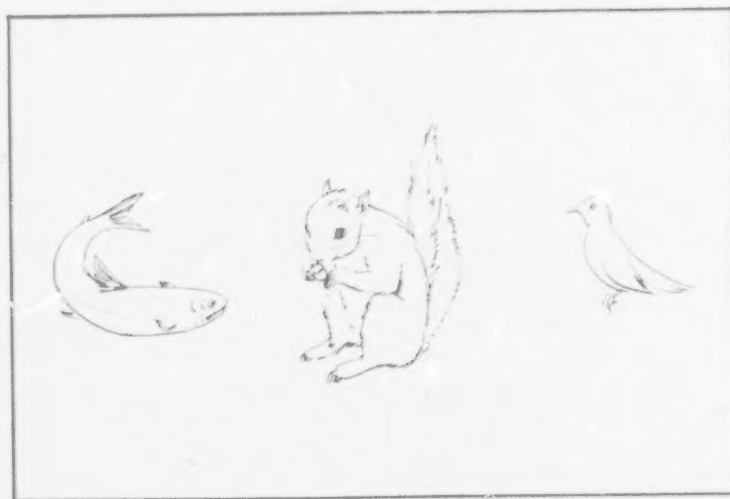
*Which one shows what most children do late at night?*

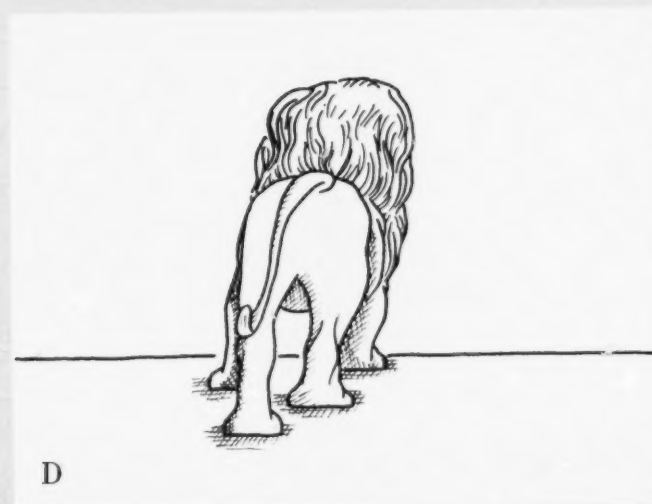
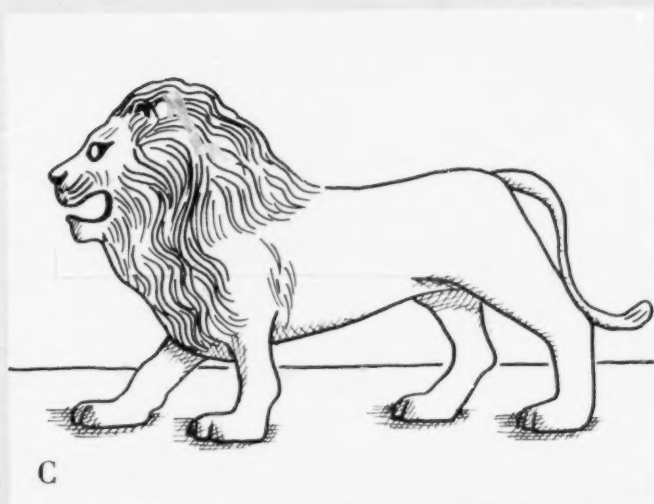
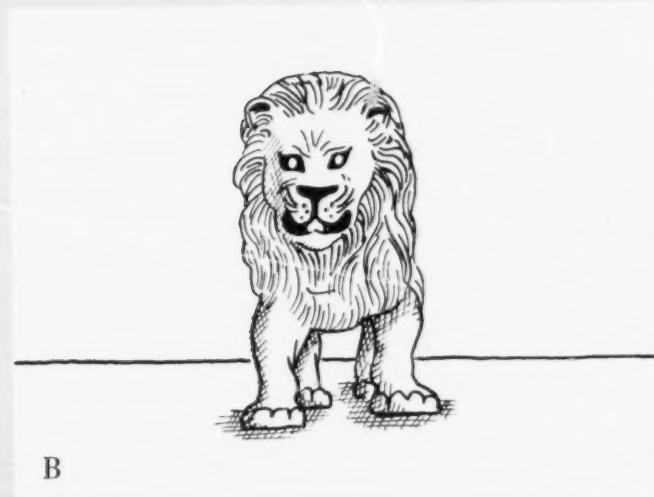
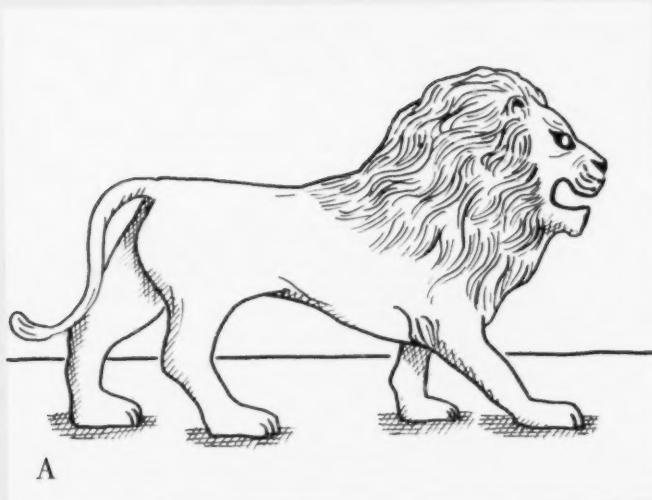


*Which one do you blow to make music?*



*Which of these has wings?*





Very young children find it difficult to understand that an object may look different when seen from a viewpoint other than their own. Using a toy lion and a toy dog, the observer places these test objects on a table at which the child is seated. Then he asks the child to indicate which of the above four cards shows how the toy lion would look from the viewpoint of the toy dog. In this way, insights are gained into the developmental phenomenon of spatial egotism. (See Anderson, *Disadvantaged Children and Their First School Experiences: A Longitudinal Study*.)



revealed favorable attitudes toward Title I programs by school superintendents, principals, teachers, and school-community representatives during both the 1967-68 and 1968-69 school years. Most wished the programs to be continued, usually with only minor changes. A mathematical model of a decision-making process whereby Title I programs might be effectively evaluated by the Chicago school system was also developed. The study explores in some depth many aspects of the evaluation of compensatory education programs. Reports are in restricted circulation.

*Objective Achievement Testing for the Evaluation of Chicago Title I Projects.*

DANIEL P. NORTON

The usefulness of standardized achievement tests for Title I project evaluation in Chicago was also the subject of a study by the ETS Midwestern Office. Specifically, use of standard achievement tests for pretesting and posttesting tens of thousands of pupils was investigated. Scores were subjected to multiple modes of analysis. Numerous artifacts—or meanings peculiar to this situation—were identified that must be understood if outcomes are to be interpreted correctly (for example, between battery differences, raw score versus grade-equivalent modes of analysis, and so forth). Reports are in restricted circulation.

*Evaluation of the Verona Plan for Sharing Educational Opportunity.*

STANLEY M. ZDEP

Will the education of children from predominantly black ghetto schools be advanced if they are bused to predominantly white suburban schools? A research study on this problem was carried out by ETS for the town of Verona, a suburb of Newark, New Jersey, to which 26 first and second graders were bused from Newark for schooling. The research design provided for comparison of the achievement of bused children with the achievement of a comparable group who remained in the ghetto schools. It was found that reading, math, and listening achievement test scores of bused first graders were significantly

higher than those of unbused first graders. No significant differences in these subject areas were found between bused second graders and their counterparts who remained in ghetto schools. The research should be of interest to educators, psychologists, and others contemplating new busing ventures, or assessing those in existence.

*Miami Bilingual Student Study.*

ROBERT F. BOLDT

The special needs of students who must live and perform in two cultures are gaining more recognition from educators. One such group consists of the bilingual Spanish- and English-speaking students, primarily of Cuban descent, who live in Miami, Florida. A practical question for the Dade County, Florida, school authorities was whether special weightings might be useful for equating the scores of bilingual students with nonbilingual students on the Scholastic Aptitude Test (SAT) and the Prueba de Aptitud Académica. This study attempted to find some basis for differential weightings using language orientation and socioeconomic ratings. It was found that use of the mathematical part of the SAT alone, or the SAT-M as it is known, seemed to be as practical and valid as any of the more complicated prediction schemes that were examined by the study. The research should be of interest to educators and test developers concerned with testing bilingual populations. CEEB. RDR-68-9. SR-69-31.

*Vocational Development Study—Patterns of Negro-White Differences.*

MICHAEL ROSENFELD

Few longitudinal studies of black and white patterns of intellectual growth have been made. Most studies of this nature are cross-sectional, with growth patterns inferred over differing sample populations. This study was unusual in observing changes in the same subjects over a period of six years. The subjects were 316 blacks and 501 whites in schools of differing curricular emphasis and ability range within the systems of two cities, one midwestern and one western. Multivariate analyses of variance were carried out on the

scores of these students at grades 5, 7, 9, and 11 on the Sequential Tests of Educational Progress (STEP) and the School and College Ability Tests (SCAT). A measure of socio-economic status was also obtained from the Background and Experience Questionnaire for these students. In general, the results corroborated the findings of most other studies that describe the intellectual performance patterns prevailing in our schools today. In general, test scores of white subjects were significantly higher than those of blacks at grade 5 on all tests of SCAT and STEP. Moreover, the whites not only started out higher, but they grew at a faster rate than the blacks. In other words, the gap widened between students of the two races from grade to grade as they progressed through high school. Such findings have been used by racists and others to support a contention of innate intellectual differences between the races. The more environmentally oriented, however, claim such findings reflect the depressing effect of socially and individually low expectancies for blacks and of possibly substandard schools and communities. Though this study by design could do little beyond record existing general performance patterns, in keeping with the environmentalist view it did find evidence of differences between schools in gains made by both blacks and whites. It also found that both blacks and whites continued to grow on all tests of SCAT and STEP, and that there was no test on which either blacks or whites showed a decline from one testing period to the next.

*Assessment of Disadvantaged Adolescents: A Different Approach to Research and Evaluation Measures.*

NORMAN E. FREEBERG

A great deal of attention has been given to the question of whether racial bias exists in many standardized tests whose use affects minorities. Attention has also been given to modifying these tests and eliminating bias. However, a surprisingly small investment has been made in the development of new measures specifically tailored to the needs, interests, and skills of these minorities. This is the purpose of this study—the

development of a battery of 13 measures specifically designed to assess the abilities of adolescents from culturally deprived, poverty-level backgrounds. Measures now in use with such groups have been heavily criticized for limitations such as entirely verbal formats, reading levels that are too high, a culturally biased content expressed in terms of majority group language and concepts, too great a length—50 to 100 items per test, being designed for large-group administration with the examiner remote from the respondent, tight time restrictions, and the like. To counter such shortcomings, the measures developed by this study are in small-booklet form, with numerous pictures, and are designed for small group administration. The verbal content is maintained at a fifth grade reading level; subject matter is relevant to disadvantaged backgrounds; and the number of items per booklet range from a minimum of 10 to a maximum of 30. The approach has been shown to have reasonable promise after having been tried with a group of 256 high school dropouts who are enrolled in 11 Neighborhood Youth Corps project centers primarily in urban areas of the northeast. Bureau of Work Training Programs, U. S. Department of Labor. RB-69-41. *Journal of Educational Psychology*—In press.

#### *Project Access Evaluation.*

AMIEL T. SHARON

Project Access is a program that is intended to help disadvantaged high school students gain access to college or other institutions of postsecondary education. This study was an evaluation of the project's potential in terms of three questions: Will students who wouldn't normally do so apply or enroll in college because of Project Access? Which parts of the projected project are effective, and which ineffective? How might the project be improved? Approximately 70,000 juniors and seniors in 99 inner-city high schools were tested with the Preliminary Scholastic Aptitude Test and completed a short questionnaire. This information was distributed to the participating high schools and to more than 500 colleges. Findings were then based on follow-up questionnaires and

interviews. Among the high schools and colleges an overwhelming majority indicated that the instruments used in the project provided them with useful information, but college personnel also offered many suggestions for improvements. Colleges experienced several problems: they had difficulty in tracking down prospective students because their home addresses were not given, were overwhelmed by the amount of data, and lacked the staff and the time to contact students. Most of the students who were interviewed had favorable reactions to Project Access. About 65 percent of them indicated that they had applied or were planning to apply to a college, while three out of every ten students said that the project made them start thinking about college. The project appears to have had a greater effect on juniors than on seniors, so its full impact will probably not be felt until fall 1970. CEEB.

#### *Dimensions of Social Stratification for Whites and Negroes.*

LAWRENCE J. STRICKER

Starting with the ground-breaking work of Lloyd Warner in "Yankee City," a vast amount of research has been devoted to developing indexes of social class that can be used to classify people as "upper class," "middle class," "lower class," and so on. Although this research has been based almost entirely on samples of whites, these indexes are often applied to Negroes, despite evidence that the indexes may not be comparable for the two groups. This study seeks, first, to identify the dimensions of social stratification for both races. Unlike previous studies, it uses samples of *both* races. It also seeks to determine whether distinct social classes actually exist—that is, whether the different classes can be clearly demarcated, or whether they simply represent gradations on a continuum, one group blending imperceptibly into another. The data are being gathered in highly structured interviews with cross-sections of both races. When completed, the research should be of interest to sociologists, psychologists, and others (social workers, teachers, educational planners) who use social class indexes in their work. Aside from

its implications for social stratification, it should also provide an interesting description of white and Negro cultures. NICHHD.

#### *Cross-Cultural Study of the Effect of Instructions on the Müller-Lyer Illusion*

JULIA CARLSON

An interesting finding is that, when confronted with a certain type of optical illusion, black Africans tend to be less susceptible to being "fooled" by it than white Americans. Segall, Campbell, and Herskovits in *The Influence of Culture on Visual Perception* attempted to explain this difference in terms of a theory of how environmental influences might account for it. By certain refinements in testing procedures, this study attempted to go beyond the Segall, Campbell, and Herskovits conclusions by relating this perceptual difference to the highly specific experimental work and theory of selective attending. The study demonstrates that selective attention is an important determinant of differences in perception, but that it will not account for the differences found by Segall, Campbell, and Herskovits. Regarding the question of innate versus environmental differences, it demonstrates how difficult it is to separate the two in research, and to obtain clear-cut findings. The research should be of interest to psychologists interested in perception and cross-cultural studies. RB-68-53.

#### *Home Environment and Educational Development.*

THOMAS L. HILTON, DONALD L. LANTZ, CATHLEEN PATRICK

Recent studies have indicated that process variables in the home—such as encouragement to continue education, parental interest in achievement, and availability of reference material—may be useful predictors of academic status. This study uses Growth Study data in an attempt to see what sort of relationship may exist between students' perceptions of process variables of this type and their actual academic growth. Academic growth was gauged by measures such as STEP and SCAT, and the students' perceptions were derived from the Test of General Information



(TGI) and Background and Experience Questionnaire (BEQ) data. A scale made up of BEQ items that seemed to relate to students' perceptions of family environment was developed. Students' scores on this scale were compared with the measures of academic growth and some interesting correlations emerged. Currently, the scale is being revised. The research should be of interest to researchers studying the relation between home environment and education.

#### *Socioeconomic Status and Academic Gain.*

THOMAS L. HILTON, CATHY JEAN COBES

This study made use of Growth Study data to explore some possible relations of socioeconomic status (SES) to academic gains. Three hypotheses for testing were derived from the work of a leading theorist in this area, Benjamin Bloom.

Hypothesis I was that gains in achievement would vary according to the SES of the student's parents. Hypothesis II was that gains in achievement over a period of time would be independent of the level of achievement at the beginning of this period—that is, that those scoring high at the beginning of the time period would show about the same rate of academic gain as those scoring low. Hypothesis III was that students grouped heterogeneously according to SES would show a significantly greater range of gain scores in achievement than students who were grouped homogeneously. Gains in achievement and mathematics were examined for a sample of 1,400 students from grades 5 through 9. No support was found for hypothesis I. Hypothesis II was rejected for mathematics, but seemed to hold for reading. No generalizable results were found for hypothesis III. The research should be of interest to re-

searchers examining the relation of SES to education.

#### *Professional/Technical Assistance to the Head Start Research and Evaluation Demonstration Program*

RUTH B. EKSTROM

The Head Start program to ready disadvantaged youngsters for school has proven to be one of the most popular social measures to emerge out of the "Great Society" program years. Part of this particular project has involved a review of a number of earlier reports on Head Start. It has also involved a classroom description study, a review of research proposals, and participation in six research seminars, the proceedings of which are now available from the Head Start office in Washington. A collection of tests for use with preschool children is also being compiled and indexed for use by the Head Start office staff. Project Head Start.

## *Career and Vocation*

#### *Bayesian Guidance Technology.*

MELVIN R. NOVICK, PAUL H. JACKSON

Much of the application of statistical theory is concerned with methods of estimating probabilities. For example, predicting a student's college performance from high school tests and grades is, at the heart of it, a matter of arriving at a judgment based on probabilities. Thus, questions of probability methods and theory are fundamental to testing. For a number of years a controversy has existed among theoreticians over the suitability or unsuitability of what are known as Bayesian methods, after the British clergyman who pioneered this approach to probability theory during the eighteenth century. Bayesian methods have been relegated to the "back shelf" in testing until quite recently because of certain unresolved complexities. However, a crucially important new development in testing is now forcing serious consideration of a fairly general imple-

mentation or switch-over to Bayesian methods in testing. This development is the shift in thinking about testing that is coming about through the widening confrontation of education with social need. That is, the traditional emphasis in testing has been on the selection of relatively few people out of many, and on a basis of a few abilities out of many, for special training. However, an increasingly well-articulated situation of social need clearly calls for more use of testing for guidance and placement of the many rather than selection of the few, and according to an evaluation of many rather than few abilities. This view is best explained within the Educational Testing Service context in an article, "Relevance in Testing," by Dr. William W. Turnbull, President of ETS. Some Bayesian theorists see their methods as ideally suited to effect the least troublesome and most effective changeover in the methods and statistical thinking necessary for broadening testing to

serve broader social ends. The issues, theory, and range of applications for test developers and statisticians are outlined in RB-69-83 by Novick and Jackson, which will appear in the October 1970 issue of the *Review of Educational Research*. NICHHD.

#### *System of Interactive Guidance and Information.*

MARTIN R. KATZ, WILLIAM F. GODWIN

A computerized, interactive guidance and information system is being developed to improve career decision making by junior college students. The project is a response to the nationwide need for career guidance for students in junior colleges. Students will be able to use the computer and related technologies to explore their own values, to obtain and interpret information, to make predictions, and to formulate and rehearse career plans. The system emphasizes wisdom in the process of career decision making rather



In-basket test flexibility allows use of a job-simulation approach in a wide range of occupational settings. The sample test materials on this and the next page are from the Foreman's Shop Desk In-Basket developed by ETS in collaboration with the Caterpillar Tractor Company. The prospective shop foreman who is being tested receives a note from another foreman, who has left in a hurry for the hospital with his pregnant wife, leaving him to carry out the work indicated on the shop line-up sheet. The prospective shop foreman must respond to this task with various work directives and assignments, on the basis of which his performance is rated. (See Crooks, *In-Basket Development and Research*.)

Pete:

I had to leave early - the wife called. Said her pains started. So I won't get to see you this A.M. Sure hope the boys clean up good at end of shift. Give 'em hell for me if they don't.

Wish me luck for a boy!

Stu

1-1

# LINE UP SHEET

SECTION 366

DATE 5-3-67

| Machine<br>Number   | Operate<br>Following<br>Shifts |              |              | Part<br>No.        |                                                                     |
|---------------------|--------------------------------|--------------|--------------|--------------------|---------------------------------------------------------------------|
|                     | 1                              | 2            | 3            |                    |                                                                     |
| 8433-34             | ✓                              | ✓            | ✓            | 5M-593             |                                                                     |
| <del>13581</del>    | <del>✓</del>                   | <del>✓</del> |              | <del>25-622</del>  | 3H-9508                                                             |
| 12714               | ✓                              | ✓            |              | 4B-870             |                                                                     |
| 14732               | ✓                              | ✓            |              | 4L-1282            |                                                                     |
| 12067               | ✓                              | ✓            |              | 3K-1589            |                                                                     |
| <del>17151-52</del> | <del>✓</del>                   | <del>✓</del> | <del>✓</del> | <del>4B-2628</del> | 2K-5634                                                             |
| 9826                | ✓                              | ✓            |              | 5H-3276            |                                                                     |
| 11427               | ✓                              | ✓            | ✓            | 4H-3485            |                                                                     |
| 10641               | ✓                              | ✓            |              | 6L-3507            |                                                                     |
| 12896               | ✓                              | ✓            |              | 5M-3520            |                                                                     |
| 8437                | ✓                              |              |              | 4F-4135            |                                                                     |
| 13138               | ✓                              |              |              | 2A-4293            |                                                                     |
| 16610               | ✓                              | ✓            |              | 9H-4889            |                                                                     |
| 10902               | ✓                              | ✓            | ✓            | 5M-5822            |                                                                     |
| W3188               | ✓                              |              |              | 2M-6548            |                                                                     |
| 7782                | ✓                              | ✓            | ✓            | 2M-6569            |                                                                     |
| 7918                | ✓                              | ✓            |              | 9M-8282            |                                                                     |
| 10914               | ✓                              | ✓            | ✓            | 4H-8576            |                                                                     |
| <del>8359</del>     | <del>✓</del>                   |              |              | <del>4H-8576</del> | 9M-8282                                                             |
| W3177               | ✓                              |              |              | 4H-8576            | 2M-6548                                                             |
| 13802               | ✓                              |              |              | 2J-6405            | 4H-8576                                                             |
| 11426               |                                |              | ✓            | 4L-1282            | ← → Steer - we need 100 by P.M.                                     |
| 14441               |                                |              |              | 3S-8800            | → Start as soon as available<br>since parts held up in<br>Loc. 2070 |

than "correctness" in conforming to some predetermined standard. Pre-eminent is the role of each student's own values. In short, the system is based on a humanistic ethic, with control vested in the student. The research and developmental work should be of interest to junior college students, guidance specialists, and junior college administrators. It also relates to research in guidance theory, vocational development, computer technology, manpower planning, and occupational information. Carnegie Corporation. Articles on the system have appeared in the *College Board Review*, Summer 1969, and in *Computer-Based Vocational Guidance Systems* (U. S. Office of Education, 1969, pp. 10-15 and 72-76).

#### *A Study of Intellectual Growth and Vocational Development.*

THOMAS L. HILTON, PATRICIA L. CASSERLY, FRANKLIN R. EVANS, NORMAN E. FREEBERG, WILLIAM F. GODWIN, MARTIN R. KATZ, MICHAEL J. PATTON, DONALD A. ROCK, MICHAEL ROSENFELD, JONATHAN R. WARREN

In this study the Growth Study data are used to trace the intellectual development of students who choose vocational rather than college preparatory curricula in high school; to examine the interaction over time of their individual natures, plans, environmental influences, and eventual occupations; and to develop a theoretical model of vocational development. A number of studies have been launched for these purposes. A Vocational Decision-Making Study developed an interview schedule to collect data on this topic. A Curriculum Group Achievement Study involved multivariate analysis of such variables as school, curriculum, sex, and socioeconomic status. In Simulation of Curriculum Assignment Process an attempt is being made to develop a computer simulation of the process whereby students become involved with a particular curriculum. In the Follow-Up Study, questionnaires were sent to graduates of Growth Study schools to gather data to be used in comparing high school progress with eventual job or school choices. The School Setting Study assessed the

effects of the communities as well as the schools on educational "climate." The Biographical Factor Stability Study determined the usefulness of the Background and Experience Questionnaire for the overall study. Moderated Prediction Within Curriculum examined extensive data on 900 students to describe how personal characteristics interact with achievement and curriculum. The Dropout Study sought to identify behavior patterns of students who drop out of high school. The Study of Negro-White Differences compared growth patterns as measured by the School and College Ability Tests and Sequential Tests of Educational Progress for Negro versus white students over grades 5, 7, 9, and 11. Of many findings, three generalizations were these: (1) though some trends were found, it was difficult to generalize, each school tending to be unique, (2) family status seems to have great impact on choice of curriculum and performance, and (3) large-scale multivariate analytical techniques, while powerful, must be supplemented with other data collection and analytical methods. USOE. RB-69-29, RB-69-32.

#### *Vocational Development Follow-Up Study.*

JONATHAN R. WARREN

The traditional view of education is that its purpose is to prepare us for what comes after we leave school—for a job, a career, or an enjoyable retirement. Thus, it is rather ironic that while numerous studies have been made of what happens to us while we're in school, relatively few follow-up studies are carried out to see what happens afterward. This study is particularly rare in being an attempt to relate the high school experience of students in all curricula—vocational as well as college-preparatory—to their post-high school experience. Questionnaires were sent to 6,800 graduates of Growth Study schools, and data for the Oakland, California, school system were extensively explored. In regard to the movement of high school graduates into jobs and colleges, one fairly general finding is quite thought-provoking. Except among girls who were in voca-

tional curricula in high school, the association between high school curriculum and post-high school activity was very weak. That is, whether or not one had taken the vocational or the college preparatory curricula seemed to show relatively little relation to whether or not one was employed or in college after high school. This suggests that differentiation of high school curricula according to whether or not a student plans to attend college may not be a wise procedure. As a whole, the study found that college attendance, as opposed to employment one year after high school graduation, was chiefly affected by the student's verbal aptitude and the social class of his parents as well as the curriculum followed in high school. Social class has a stronger effect on high school curriculum choice for women than for men. However, aptitude seems to be a stronger influence than social class for both men and women. USOE.

#### *Effects of Organizational Climates.*

NORMAN FREDERIKSEN, OLLIE JENSEN, ALBERT E. BEATON, BRUCE BLOXOM.

There is a tendency among personnel workers to believe that performance can be predicted from ability and personality measures, without taking much account of variations in work situations. This is a questionable belief, since situational variables may interact with personal characteristics to affect performance. The present study explored this general question area in considerable depth with data from 260 executives, each of whom served for two days as chief of a hypothetical field services division of a state department of commerce, using an in-basket test that simulated the paper work of an executive. Half of the subjects were led to believe that the department of commerce encouraged imaginative solutions to problems and cared little about rules, whereas others were given the impression that it required close adherence to rules and regulations and discouraged innovation. An analysis of responses to the in-basket problem revealed 10 aspects of performance that could be considered in evaluating an executive, such as his productivity, tendency to analyze problems, pro-



crastination, and informality. The interrelationships of these criteria were influenced by organizational climates. For example, thoughtful analysis was associated with a tendency to work through one's own organization in the rules climate, whereas in the innovation climate thoughtful analysis was associated with efforts to get out in the field and deal directly with the source of the problems. The research should be of interest to personnel officers and administrators of large organizations and to researchers in personality and social behavior. ONR, California State Personnel Board. RB-68-41.

#### *The Executive Study.*

LOIS A. CROOKS

This is a project initiated by ETS in 1955 to stimulate research on various aspects of selecting, assessing, and developing executives. A number of companies joined together to support and participate in research of this nature, and in 1964 organized formally as the Executive Study Conference. Now comprising 25 companies, the group meets twice annually to discuss research related to this area. Reports of the meetings are published and should be of interest to personnel managers and industrial research specialists.

#### *In-Basket Development and Research.*

LOIS A. CROOKS

The In-Basket Test is an approach to evaluation that attempts to assess managerial potential by simulating the problem-solving and administrative functions of an executive in terms of the flow of paperwork in his in-basket. That is, certain incoming memoranda present a problem, he records his decisions regarding it in outgoing memoranda—and these outgoing memoranda are then evaluated according to standards developed for each specific use of the technique. For several years ETS has carried out research to develop the technique and apply it to specific organizational needs. For example, General Electric has administered a Plant Manager In-Basket Test, developed in cooperation with ETS, to more than 1,200 managers, with ETS scoring the

results and modifying the test as needed. The Consolidated Fund In-Basket—the only exercise of this type currently available at ETS for purchase and general use—has the subject assume the role of a paid community fund director. Currently the subject of follow-up studies is another in-basket test developed in cooperation with the J. C. Penney Company for use in assessing potential service-center managers. The research should be of interest to those interested in managerial assessment and simulated task testing.

#### *Construct Validity of the CLEP Educational Psychology Subject Examination.*

ROBERTA B. FLESHER

The College-Level Examination Program, or CLEP, as it is generally known, is a successful new attempt to answer a long-recognized educational need. For years, the self-educated—or those well-educated in spite of poor schools—have wondered why there couldn't be some test they might take for which they might receive credit equivalent to that normally given after formal coursework at recognized institutions. Now there is such a program; CLEP offers tests of this nature in several fields. This study was to determine the validity of a new CLEP test of proficiency as a result of instruction in an introductory educational psychology course. Two large state universities participated in the validity study by administering the new examination at the beginning and at the end of their introductory educational psychology courses. A significant increase in the average score from pretest to posttest indicated that the new test is a valid measure. The research should be of interest to test developers. SR-69-9. CEEB.

#### *WGBH-TV Training Program for Law Enforcement Officials.*

RODERICK IRONSIDE

Television's potential as an instructional medium is highly attractive. However, it poses many problems not only

of programming, but also of *evaluating* programming. This project attempted to assess the impact of a series of five one-hour telecasts beamed to police officers throughout New England. The programs were concerned with police responsibilities and suspects' rights, as affected by the 1964-67 Supreme Court rulings. Prior to the programs, tests were mailed to a random sample of officers. Though response was meager and apparently selective, those who did respond indicated favorable reactions to the program content and format. However, in a pretest-posttest comparison, no meaningful difference was found between those who had and those who had not seen the programs. The study should be of interest to researchers concerned with the problems of random sampling, test development, and teaching by television.

#### *Differential Behavior Patterns.*

GEORGE G. GORDON

In obtaining names of prospective customers to call on, insurance salesmen have a choice of personal sources (such as referrals from past customers) versus impersonal sources (such as the telephone book). In training salesmen, many insurance companies stress the use of personal sources as the best method for all salesmen to use. This study, however, explored the question of whether one method was best for all. The results of personality tests given to 435 salesmen were compared with their managers' descriptions of their actual performance on the job. It was found that while men with high ego strength did do best with leads from personal sources, those with low ego strength did not. A conclusion was that it is a mistake to base training entirely on the so-called top or ideal producer when the actual operation of insurance selling calls for some variation in types of men and the approaches natural to them. The study should be of interest to the insurance industry and other industries employing salesmen. An educational implication is that training should fit the man and the situation, rather than some logically derived "ideal."

## Evaluation of Educational Systems

### *National Assessment of Educational Progress.*

SCARVIA B. ANDERSON,  
ARLEEN S. BARRON

What is the status of education in the United States? Though the educational progress of the individual student can be assessed from grade to grade, there is at present no system for assessing the effectiveness of education as a whole throughout the country, or for determining whether overall educational progress is being made. Educational Testing Service is developing measures to be used for regular national assessment of nationwide samples. The first administration of tests in three areas, science and writing (both by ETS) and citizenship (by AIR) has been completed. Results may be available in fall 1970. Pretests and special studies have already indicated some aspects of particular interest—for example, it has been found that school-age children not only can handle, but in fact enjoy, unusual test formats, such as the use of an "I don't know" option added to the standard right-wrong, true-false format, the use of paced-tape presentation of questions, and "test packages" made up of exercises from several areas and using a variety of formats. The sample for this study is of four ages (9, 13, 17, and young adults), both sexes, four regions, four community sizes, and two economic conditions; race may also be included as a variable. Two or three content areas are included in each assessment year; reading and literature will be assessed in 1970-71. The research should be of interest to curriculum specialists, educators, and school administrators, and laymen concerned with education. In addition to its development of useful measures, it will provide a large data bank for researchers and for test and curriculum developers.

### *Educational Systems Information Program.*

DONALD A. TRISMEN, GARY L. MARCO

In this time of rapid social change

many school administrators are confronted with a dilemma in planning and in accountability. They face such questions as "What should be the goals for my particular school or school system?" And "How and where should the money be spent and how can I justify it?" These two questions are tied together, for if one can specify clear-cut goals, and then be able to show progress toward them, one has solved the nub of the problems of both effective planning and cost accountability. The Educational Systems Information Program, or ESIP, is being designed to provide such a service to administrators and school boards. Using special sampling techniques, ESIP would provide an accurate, comprehensive picture of a school in terms of the nature of its students, the homes they come from, the part of the community the school serves, and a range of student qualities and skills far beyond the customary important but narrow academic ability range. As projected, ESIP would present this information in a way that would make it possible for an administrator to periodically assess progress toward specific educational goals. Consultations with interested school authorities and pilot studies are under way. In Cincinnati, Ohio, techniques of item sampling in the areas of mathematics in grades 8 and 9, and socioeconomic problems in grade 12 are being studied. The project should be of interest to community leaders as well as school administrators and boards of education.

### *Methods of Measuring School System Performance.*

HENRY S. DYER, ROBERT L. LINN,  
MICHAEL J. PATTON

For years the emphasis in testing has been upon evaluating the individual in terms of how well he measures up to the requirements of his school. However, in keeping with other trends in this direction a new interest is developing in evaluating the school and the school system to see how well they are doing by the students. That is the purpose of this study, and in some regards

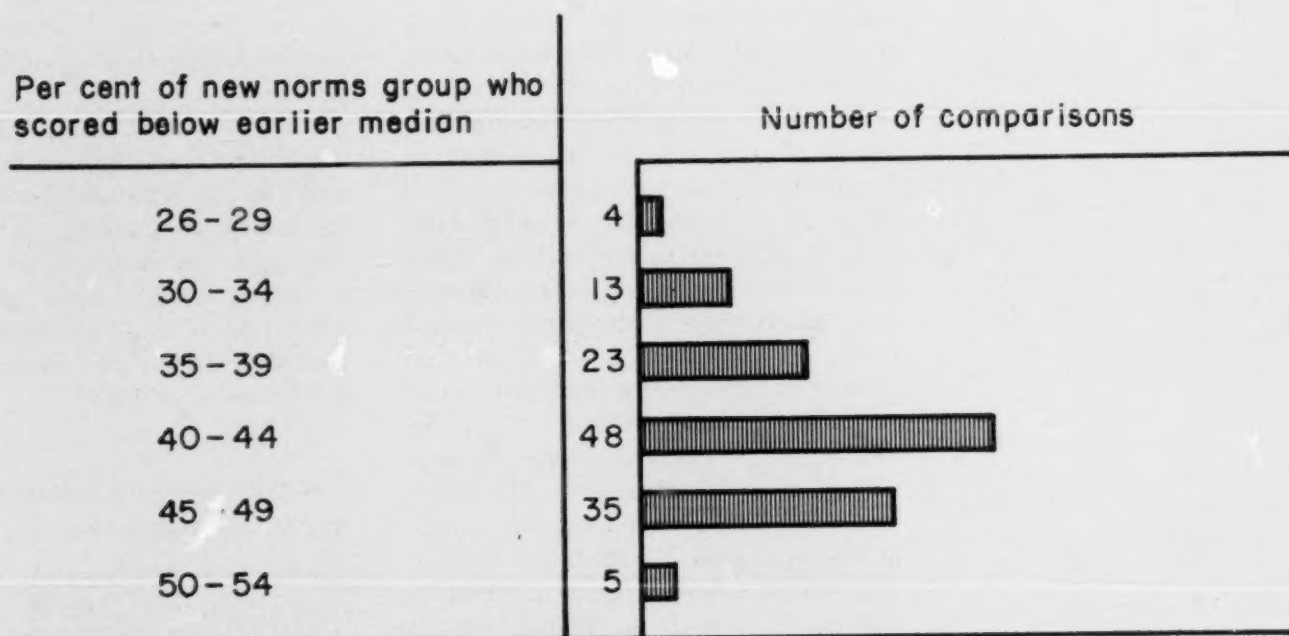
it is a pioneering study of an extremely difficult area. The basic approach used is to obtain regression lines for a measure of student performance at one grade level compared with student performance measures at an earlier grade level. Deviations of school-system means from these regression surfaces are then used as indicators of school-system performance. Using this approach, the possibilities of four different types of data were compared: (1) matched longitudinal data using individual student scores for the regression analysis, (2) matched longitudinal data using system means, (3) unmatched longitudinal data using system means, and (4) cross-sectional data using system means. The findings indicated that the first two methods yielded highly similar results and show promise. The latter two methods yielded results dissimilar from one another, and from the first two methods, and are not reasonable substitutes for the first two methods. The research should be of interest to educational researchers and educational systems administrators and planners. RB-68-55. *American Educational Research Journal*, 6:4 (November 1969).

### *Test Data as Social Indicators.*

WILLIAM B. SCHRADER

Are we being educated any better today than we were, say, 10 years ago? Based on expenditures for education one would assume that we are. But knowing for sure would depend on isolating some dependable indicators. This study attempted to do so by carefully examining the body of work most relevant to this question—the extensive educational testing activities in the United States. Could any shift in performance level be detected over the past 10 years? It was judged that national norms developed for six widely used tests were sufficiently comparable to throw light on possible trends. Examination of data on high school students provided no clear-cut answer one way or the other. However, three independent norms studies, which in-

Is American education getting any better over the decades? Answers to this question are usually a matter of opinion, with the assessments ranging from great optimism to predictions of scholastic bankruptcy. This chart summarizes some data that support optimism. Norms based on the testing of various groups in the 1950s were compared with norms for the same tests and comparable groups in the 1960s. In order to compare the groups the percentage of students in the new norms group who would score at or below the median of the old group was calculated. Thus, the smaller the percentage scoring below the old median, the greater the improvement shown by the new group. As the chart indicates, in all but five of the comparisons the new group was superior to the old group in performance. (See Schrader, *Test Data as Social Indicators*.)



Results of 128 comparisons of performance of earlier (1952-1958) and later (1963-1966) norms groups composed of students in Grades 5 through 8



cluded data for grades 5 through 8, indicated that for this age group performance during the 1960s was somewhat better than performance during the 1950s. The research should be of interest to those interested in systems evaluations of education. USOE. SR-68-77.

#### *Validation Plans for Driver Education Training Programs.*

HARRY H. HARMAN

The National Highway Safety Bureau (NHSB) is confronted with a sizable new responsibility. Federal funds have been specifically earmarked for distribution to the states for use in programs aimed at cutting the rate of deaths and accidents on our highways. One countermeasure proposed was improved driver training. The NHSB funded four studies to determine how states might best go about setting up such programs. Then it turned to ETS for a final systems evaluation, including the findings of the four previous studies. That is, by viewing the operation of driver training programs as a system involving complex interactions ETS developed a plan for making scientific evaluations of these programs. In addition to the ETS staff, a group of nationally known experts provided advice throughout the course of the project. On the basis of the project report, the NHSB is now implementing plans for

evaluating driver education and training programs. The research should be of interest to those concerned with similar problems of practical, short-term education, as well as those interested in the application of systems analysis to social problems. PR-68-7, PR-69-1.

#### *Evaluation of a Summer Reading Institute.*

MICHAEL ROSENFELD

During the summer of 1968, the Model School Division (MSD) of Washington, D. C. conducted an institute in beginning reading for teachers in kindergarten through the third grade. The purpose of the institute was to improve reading instruction by training teachers in several methods. Educational Testing Service was asked to help MSD personnel develop formal objectives for the institute, a design for evaluative research, and tests tailored to the research need. After developing objectives, a questionnaire was prepared and administered to 69 teachers at the end of the six-week training session. Questionnaire data indicated a majority of participants felt they had learned quite a bit about methods of teaching reading, felt the institute was relevant to their needs, planned to put what they had learned into action, and generally liked the way the institute

was conducted. The research should be of interest to teachers and researchers in reading instruction.

#### *Matrix Item Sampling.*

LEONARD S. CAHEN

Both tester and testee are becoming sensitive to the dangers of overtesting. One promising technique for using available time more efficiently is known as item-sampling. For example, group means of test scores are customarily obtained by processing the score totals of all individuals in whatever the group may be—a class, a school, a subject-matter grouping across the nation. Valid group means, however, may also be obtained either from the scores of a limited sample of the total number of students or by sampling the responses to single items within the total test. This study is exploring item-sampling techniques with high school students. Indications are that the methods being used are efficient for estimating school means for ninth grade students, and the approach is also being tried with twelfth grade students. An account of this study, co-authored with Thomas A. Temberg and Walter Zwirner, will appear in *Educational and Psychological Measurement*, Spring, 1970. The research should be of interest to educators, educational researchers, and evaluators. NICHHD.

### III

---

## THEORIES, METHODS, AND SYSTEMS RESEARCH

## Psychometric Theory

### *Strong True-Score Theory.*

FREDERIC M. LORD, MARILYN S. WINGERSKY, MARTHA STOCKING

How to make inferences about an examinee's ability from his test scores is a traditional, persisting problem that calls for persisting research. This study found that a mathematical model for relating what is known as the true score to the observed score fits the data very well and successfully predicts observable phenomena. The model can also be used to fit observed-score frequency distributions. It assumes that for a fixed true score, the observed score has a compound binomial distribution. The research should be of interest to psychometricians and statisticians. Its use in education is for obtaining test norms when the entire test cannot be administered to a nationally representative sample. ONR. RB-68-8. RM-69-2. RM-69-4.

### *Project Mirkwood.*

FREDERIC M. LORD, MARTHA STOCKING

A method of testing that has been arousing wide interest involves the restriction of guessing. For example, in this project the examinee is told not to guess unless he thinks he has a 50-50 chance of being right. The goal is to see whether it is possible to increase the amount of information obtained from a multiple-choice test by restricting guessing in this fashion. Two findings were: (1) theoretical computations show that this is indeed possible, but (2) in practice, the actual results did not yield better measurement. The research should be of interest to mental test theorists and test constructors, and is relevant to the study of scoring procedures based on confidence weightings. As for the title, fans of J. R. R. Tolkien will recognize a familiar name: Mirkwood. Searching for a meaningful relation to the project itself, however, will prove a fruitless guessing game. The fact is Dr. Lord and Miss Stocking happened to be reading the same book at the same time, were confronted with

the need for a project title, and on impulse decided to commemorate their enjoyment of Tolkien.

### *Tailored Testing Theory.*

FREDERIC M. LORD, MARTHA STOCKING

The successful experimental use of computers for instruction implies that they may some day be extensively used for testing. Of the types of tests which seem promising, tailored tests would be most easily administered by computer. A tailored test is one in which an examinee's response to each question determines which of possibly several choices of questions will be given to him next in the testing sequence. A number of empirical studies of tailored tests have been made by the Army, at ETS, and elsewhere. However, the purpose of this study was to see to what extent the knowledge gained from empirical testing can be generalized through the use of theory. It found that the effectiveness of the various tailored testing procedures can actually be predicted using the appropriate mental test theory. It also found that tailored tests provide more information than standard tests about an examinee at extreme ability levels—that is, either high or low proficiency. This is particularly relevant to educational testing in that the more information that can be obtained from testing, the easier it will be to predict the probable future course of an examinee's education. The research should be of interest to mental test theorists, test constructors, and those interested in computerized testing. ONR. RB-68-38. RB-69-18. RB-69-63.

### *Statistical Theories of Mental Test Scores.*

MELVIN R. NOVICK, FREDERIC M. LORD

More than 15 years have passed since the publication of *Theory of Mental Tests* by Harold Gulliksen of ETS. Now a new synthesis of the field has been prepared by Drs. Novick and Lord. Their text, *Statistical Theories of Mental Test*

*Scores*, was published in 1968. The book gives precise statements of the assumptions underlying and implications deriving from a number of test theory models. It also attempts an integration of these models into a very general theory of latent traits. Four chapters on latent trait theory and the logistic sequence model are by Allan Birnbaum of New York University. The project's present task is to review all new work in test theory and integrate it into the structure of this text for revised editions. This research should be of interest to psychometricians and test theorists. ONR.

### *Statistical Models of Mental Tests.*

MELVIN R. NOVICK

Most people think of psychological research as being a matter of testing people or animals and drawing conclusions from what happens. These conclusions then become the raw material out of which theories are constructed. Another approach to psychological truth, however, is that of seeking certain basic, seemingly self-evident statements about reality upon which one can base theories. These basic statements are known as axioms, and the mathematical psychologist—who in addition to 100 years of scientific psychology has more than 2,500 years of scientific mathematics at his disposal—is one of few who are equipped to seek them. This study is a search for axioms to which all modern mental testing theory might be related and upon which a new general theory of testing might be constructed. Several papers in professional journals cover work on the project to date. The research should be of interest to testing theoreticians. ONR. NICHHD.

### *Statistical Methods in Personnel Selection, Evaluation, and Guidance.*

MELVIN R. NOVICK

Tests have been very useful for matching applicants' skills and interests with the needs of particular jobs, in evaluat-



ing an employee's performance on the job, as an aid in job counseling, and in selecting students. A major synthesis of selection methods bearing on this area was accomplished by R. L. Thorndike soon after World War II. Since then, however, new developments in prediction methods have brought on the need for a new synthesis of the underlying testing theory. But before this can happen, certain problems must be solved. One is the problem of correcting for a limited sample—that is, when only a small percentage of a group is selected for testing, the findings will produce rather grave distortions unless one uses dependable means of statistical manipulation before using the data to predict the nature or actions of the group as a whole. Another problem is to resolve certain statistical difficulties involved in selecting the right group of variables for prediction in future samples. Publications covering several phases of the work to date are available through ETS and other sources. The research should be of interest to those concerned with test theory and developing computer programs that will facilitate selecting, evaluating, and counseling students and employees. ONR.

#### *Bayesian Inference.*

MELVIN R. NOVICK

A satisfactory solution to the long-standing problem of specifying prior distributions over the states of nature required for Bayesian analysis is needed. The purpose of this research is to seek such a solution and also to explicate a dual credibility-frequency theory of probability. To date, the prior distributions derived from the proposed indifference procedures have been found to be in substantial agreement with some of those previously accepted on essentially intuitive bases. However, the present system is often able to "explain" differences that have arisen among earlier specifications. The research is covered in detail in the *Journal of the American Statistical Association* (December and March, 1965) and the *Journal of the Royal Statistical Society* (1968). NICHHD.

#### *A General Method for Analysis of Covariance Structures.*

KARL JORESOKG

Selecting and tailoring statistical techniques to fit the peculiarities of various practical problems can be not only time-consuming but also a bewildering process. Hence, seeking generally applicable solutions for a multitude of problems is a matter of urgency in research. This study sought to develop a statistical model, and an associated method of data analysis, so general that it could be used to handle many different kinds of data and problems in the behavioral sciences. The most notable finding was that the technique under investigation can be successfully used to perform: 1) analysis of sets of congeneric tests; 2) first- and second-order factor analysis; 3) analysis of simplexes and circumplexes; 4) analysis of multi-method-multitrait data; 5) analysis of multi-test-multioccasion data; 6) analysis of linear structural relationships; 7) mixed and random effects ANOVA; and 8) MANOVA and other kinds of advanced analysis. Besides its generality and flexibility, the method is unusual within this sphere. The research should be of interest to data analysts in the behavioral sciences and to statisticians. A condensed report will appear in *Biometrika*. RB-69-46. RB-69-47. RB-69-62. NSF.

#### *A General Theory of Oblique Factor Rotation.*

WALTER KRISTOF

One of the most challenging aspects of factor analysis is the matter of rotation and interpretation. Within this general area a specific problem of importance is how to state the necessary and sufficient conditions for rotating factor patterns and factor structures to each other with a high degree of fit. This study states these conditions in terms of theory of eigenvectors and eigenvalues. It also found that fit can often be improved through a slight change of the target matrix. The procedure can lead to improvements in comparing results from factor analyses, and should be of interest to both theoretical and practical factor analysts. NICHHD.

#### *On Proportional Profiles in Factor Analysis.*

WALTER KRISTOF

Questions of existence, ways of transformation, and goodness of fit criteria for proportional factor profiles were explored by this study. By placing special emphasis on symmetrical goodness of fit criteria, it found that necessary and sufficient conditions for the existence of proportional factor profiles can be established. It also found ways of reducing the problem of proportional factor *patterns* to the more easily manipulated problem of proportional factor *profiles*. The study revises and complements basic work on an important aspect of factor analysis, and should be of interest to factor analysts and psychometricians. NICHHD.

#### *Statistical Notes on Reliability Estimation.*

WALTER KRISTOF

When a test has been split into any two parts under normality assumptions, obtaining a maximum-likelihood estimation of test reliability (coefficient alpha) becomes a problem. Using the maximum likelihood method, this study identifies the estimator. It also gives a statistical test of hypotheses concerning the parameter that establishes confidence intervals. It is chiefly relevant to other research in this area by generalizing and closing gaps in the research. This study research should be of interest to test theorists, test users, and statisticians. RB-69-25. NICHHD.

#### *Structural Properties and Measurement Theory of Certain Sets Admitting a Concatenation Operation.*

WALTER KRISTOF

Explorations in advanced measurement theory have raised questions of the existence and uniqueness of scales for certain continuous and ordered sets when a concatenation operation for pairs of objects is defined which satisfies certain requirements. This study found that scales of this nature can be constructed and are unique up to positive linear transformations. It also developed a number of theorems concerning reflexivity and commutativity

of the concatenation operation. The findings complement other theories, and should be of interest to measurement theorists, researchers in psychophysics, and mathematicians. *British Journal of Mathematical and Statistical Psychology*, 1968, 21, 201-229. NICHHD.

*A Theorem on the Trace of Certain Matrix Products and Some Applications.*

WALTER KRISTOF

A problem in the use of matrix theory is the stating of attainable upper and lower limits for the trace of certain matrix products. This study found a solution for the problem by stating these limits as a theorem, which was proved by induction with an emphasis on canonical decompositions of matrices. The theorem can be used to solve various problems in psychometrics when the trace of matrices is involved and should be of interest to psychometricians, mathematicians, and statisticians. RB-69-21. *Journal of Mathematical Psychology* (In press). NICHHD.

*Estimation of True Score and Error Variance for Tests under Various Equivalence Assumptions.*

WALTER KRISTOF

Another problem in the field of mental test theory is how to develop maximum-likelihood estimators of true score variance and error variance for mental tests under different hypotheses of equivalent measurements. Using a measurement model involving classical test theory in conjunction with distributional assumptions, this study established these estimators and some of their statistical properties. It also developed a statistical criterion for tau-equivalent measurements. The research should be of interest to test theorists, test users, and statisticians. RB-68-57. *Psychometrika*, December, 1969. NICHHD.

*On the Parallelization of Trace Lines for a Certain Test Model.*

WALTER KRISTOF

Whether parallelization of a set of

trace lines is possible for a certain test model is a problem for the measurement theorist. Using the theory of functional equations, this study found that parallelization is possible and essentially unique if—and only if—all item characteristic functions are powers of each other. It also found that upon transformation all item characteristic functions assume a unique functional form regardless of their initial form. The findings shed light on the notion of latent continua and on the possible proper form of characteristic functions. The research should be of interest to test theorists. RB-68-56. NICHHD.

*On the Sampling Theory of Reliability Estimation.*

WALTER KRISTOF

This study sought to develop a statistical test for coefficient alpha—and hence reliability—under the situation of normality assumptions when a mental test has been split into any two parts. While this is customarily a difficult problem, by a certain transformation of variables the problem was reduced to more workable and familiar proportions. The end results of the study were that a strict test for coefficient alpha for any sample size was developed, and confidence intervals for the parameters were established. The research should be of interest to test theorists, test users, and psychometricians. RB-69-36. Accepted for publication by the *Journal of Mathematical Psychology*. ONR.

*On Metric Multidimensional Unfolding.*

PETER SCHONEMANN

The problem of locating two sets of points in a joint space when you are given the Euclidean distances between elements from distinct sets was solved algebraically in this study. For error free data the solution is exact; for fallible data it has least squares properties. A computer program was written to carry out the necessary computations. Several numerical illustrations are given in the study to illustrate the feasibility of using this procedure for the analysis of nonmetric data. The

research should be of interest to psychometricians. *Psychometrika* (in press).

*Scaling a Simplex Symmetrically.*

PETER SCHONEMANN

The purpose of this study was to devise a method for estimating the latent parameters of simplex matrices that does not depend on the order of the manifest variables. A method for scaling a perfect simplex is suggested which is based on a procedure for scaling a set of points from their pairwise distances, thus retaining the symmetry of the data matrix. The approach was extended to an iterative algorithm for scaling a quasi-simplex, and a FORTRAN IV program for scaling both types of simplex matrices was written. Among the findings were certain indeterminacies in the quasi-simplex model, one of which does not seem to have received much attention so far, although it may constrict the practical utility of the more general model. On applying the program to a number of data sets from the literature, it was found that the fit for a quasi-simplex was usually satisfactory, while the fit for a perfect simplex was usually poor. The research should be of interest to psychometricians. RB-68-31. *Psychometrika*, March 1970.

*Fitting One Matrix to Another Under Choice of a Central Dilation and a Rigid Motion.*

PETER SCHONEMANN

The purpose of this study was to develop a least squares procedure to compare results obtained with nonmetric scaling techniques across samples, or to compare such results with those obtained by conventional factor analysis on the same sample. A least squares method was derived for fitting a given matrix A to another given matrix B under choice of an unknown rotation, an unknown translation, and an unknown central dilation. The method was programmed for the computer and applied to several data sets, both real and artificial. The research should be of interest to psychometricians. RB-69-78. *Psychometrika* (in press).



### *On Two-Sided Orthogonal Procrustes Problems.*

PETER SCHONEMANN

Earlier work of Dr. Schonemann dealt with a least squares solution to the problem of finding an orthogonal matrix  $T$  in  $B = T'AT = E$ , given  $A, B$ , both symmetric, which minimizes  $\text{tr}E'E$ . It was hoped that the solution of this problem might prove useful in solving another problem: Given  $A, B$ , both symmetric, find a permutation matrix  $P$  in  $B = P'AP + E$ , which minimizes  $\text{tr}E'E$ . One difficulty with this approach is that the solution matrix  $T$  of the continuous problem is not unique. Instead, a family of  $2p$  (if  $A, B$  are of order  $p \times p$ ) solution matrices exist, all of which reduce the loss function to the same minimum. The present study was concerned with finding practical means for removing this indeterminacy to make the suggested method useful for the intended purpose. The attempt, however, proved unsuccessful. NIMH.

### *Variance Analysis System—MANOVA.*

CHARLES E. HALL

As knowledge increases by leaps and bounds so grows the need to expand and simplify our methods of processing and making use of it. The aim of this project is to increase the capacity for statistical processing of information by developing a general system for correlation analysis, factor analysis, and analysis of variance and covariance so as to eliminate separate computer programs. In moving toward this objective, some general programs for statistical processes are being studied at ETS: special attention has been given to two multivariate analysis of variance programs: MANOVA, by Elliot Cramer, and MULTIVARIANCE, by Jeremy Finn. In addition, an improved method for computing orthogonal polynomial coefficients has been devised, and efficient methods for matrix inversion are being explored. The research should be of interest to psychometricians and statisticians working with researchers.

### *Advanced Moderator Variable Study.*

DONALD A. ROCK, FRANKLIN R. EVANS

A way of increasing the predictive validity of tests is to identify, and then

construct tests to make use of, moderators as well as predictor variables. For example, where a high score on a test of quantitative ability might to a certain degree predict math grades in college, this predictability might be increased if one knows whether the candidate has a parent who wants to see him succeed—a variable which in this hypothetical instance would moderate the usefulness of the predictor variable in estimating probable math performance. This study is investigating the moderating effects of certain background variables on the prediction of law school grades. It is also examining whether a number of new tests would act as moderators or as predictors. A measure to assess certain personality traits is being constructed. Earlier a battery of tests was assembled and pretested for final administration during the fall of 1969. Results of this testing will be used in constructing a final battery. The research is for the Law School Admission Test Council.

### *Gauss-Seidel Computing Procedures for Some Methods of Factor Analysis.*

MICHAEL BROWNE

The purpose of this study was to use Gauss-Seidel computing procedures to obtain the following estimates of the factor matrix: Type I, maximum likelihood; Type II, alpha; Type III, least squares with fixed weights; Type IA, weighted least squares using residual variances as weights; and type IIA, weighted least squares using common variances as weights. Among the findings, the study suggests modifications in Gauss-Seidel computing procedures in obtaining Type I and Type III solutions. Procedures for solving the nonlinear equations involved in Type IA and Type IIA solutions were developed. Subroutines were also prepared for each of the five solutions. The results generally indicate that the procedures are effective with a number of empirical correlation matrices, including two non-Gramian tetrachoric correlation matrices. However, the Type I solution should not be computed if the correlation matrix is non-Gramian. The research should be of interest to psychometricians and test developers. RB-68-61.

### *Precision of Prediction.*

MICHAEL BROWNE

In psychology and psychometrics, much effort goes into developing theories that can be used to predict behavioral and other outcomes. The task is crucial not only because it is useful to be able to predict what people will do under certain circumstances, but also because theories serve generally to establish firm ground within the great open reaches of psychological space. During the last 20 years, the research psychologist's powers of prediction have been tremendously increased by a new statistical and computer methodology that allows him to take into account many rather than few variables in his research explorations. This new power, however, has led him to confront an important question: How far may one go in taking predictor variables into account before one reaches a point of diminishing return? This is a question of trade-offs, perhaps more familiar as: Will one hundred dollars' worth of additional effort produce two hundred or only two dollars' worth of benefit? The question was explored mathematically and also by programming a computer to many kinds of trade-off situations. The optimum number of predictor variables was found to depend on sample size. A formula, which should aid psychologists in gaining precision of prediction, has been developed to express this relationship. NICHHD. RB-69-69.

### *Least Squares Matrix Factorization and Multidimensional Scaling.*

GORDON G. BECHTEL

A problem that might at first glance seem rather remote from everyday life is that of deriving multidimensional scales from the choices individuals make among stimuli. The major finding in this study was a theoretical point involved in the way in which the least squares matrix factorization method will provide the key to solving this problem. However, an application reported in the study begins to suggest its generality. The dimensions of "liberalism-conservatism" and "political audacity" were found when this method was applied to data collected



on political preferences. The research should be of interest to psychometricians, psychologists, economists, sociologists, market researchers, political scientists, opinion and attitude polling organizations, and some educators. It provides a method whereby researchers in the above fields can discover the dimensions underlying whatever kinds of choice behavior they may wish to study. For example, in education the study may be useful in discovering the dimensions that underlie social interaction in the classroom, or the dimensions underlying student choices among courses, vocations, other students, and so on. RB-69-73.

#### *Defining Groups in Multivariate Space*

BARY G. WINGERSKY

In arriving at many practical decisions involving large numbers of people, it is necessary to separate them into categories according to test scores. For example, test-takers can be separated into the familiar pass or fail categories according to the results of a quiz relating to a particular ability. There are far more complex situations, however, where the scores of individuals must be grouped into several categories according to responses varying along numerous personality, background, or cognitive

ability dimensions. This study has developed a way of identifying useful categories within a body of data and a procedure for manipulating the data to make these structures more evident. Moreover, a computer program for indicating whether such structures within a body of data relate to one main grouping, or to many groupings, has been completed. The procedure is being tested with personality inventory scores of 602 college boys and 679 college girls, and programs based on the procedure should eventually be available for general use. The research should be of interest to test developers and research data processors.

## *Methodological Research*

#### *Statistical Adjustments When Comparing Preexisting Groups.*

FREDERIC M. LORD

It is common to use analysis of covariance to "adjust" data so that preformed groups (e.g., racial groups) can be compared on final achievement with their initial differences supposedly removed. This study explored the use of this approach in correcting for preexisting differences between preformed groups, and found that analysis of covariance is often clearly inadequate. The research should be of interest to anyone who wishes to correct for initial differences. It is particularly relevant to comparisons of racial or economic groups on ability to learn. RB-68-67. *Psychological Bulletin*, 1969, 72, 336-337.

#### *Learning.*

HAROLD GULLIKSEN

An appealing feature of mathematical learning theory is that it makes it possible to study the learning process in terms of the precision offered by, for example, the plotting of learning task performance curves on a graph. A difficulty, however, is that there is so much variation between individuals and between groups in learning capacity

that much of the advantage of the mathematical learning theory and of curve-fitting statistical methods is lost through the imprecisions of the usual multi-organism comparisons. For this reason researchers have become greatly intrigued by the possibilities of the split-brain technique, whereby it is theoretically possible to compare learning performance curves not between two individuals or groups, but within or between the two halves of the same brain. The assumption has been that right and left brain learning is identical. This research, however, seeks to determine whether this is actually true. The parameter estimation is still proceeding, but to date it appears that right and left brain learning are not identical. One noteworthy contribution has been a new parameter estimation procedure developed by Dr. Gulliksen's student, James Ramsay. The research should be of interest to researchers concerned with mathematical learning models, split-brain work, and parameter estimation methods.

#### *Independent Activities Questionnaire.*

FRANKLIN R. EVANS

Most formal testing purposely disregards the past in order to concen-

trate on seeing what a person can actually do at one present point in time. One reason for the popularity of this approach is that it lends itself to rather clean-cut and efficient processing, and the past is not so easily quantified. However, there is much to be said for an approach that attempts to quantify past performance by viewing it as the accumulation of real life achievements. Such an instrument is the Independent Activities Questionnaire (IAQ), which quantifies reports of a great variety of high-level achievements in activities that are self-initiated by students during their high school years. The IAQ, for example, provides a way of quantifying some possibly important indicators of adult potential such as high school activities—being a class president, winning an art prize, or building a successful hot rod car. To date IAQ has been used chiefly as a criterion measure for creativity studies. However, continuing research and evaluation of the instrument indicates many other potential uses. New forms have been developed for use in research into the creative potentials of college and graduate students. The approach also seems to have developmental possibilities for assessing the potential of black students, or students from any

other cultural background where potential for "success" may be better reflected by performance in their environment than in performance on a formal test. The research should be of interest to guidance and admissions personnel at all educational levels, as well as researchers in creativity and black and urban problem studies.

#### *Longitudinal Study of Early Predictors of Creative Achievements.*

FRANKLIN R. EVANS

This study is investigating whether measures of student ability, interest, and achievement in grades 7, 9, 11, and 12 can predict IAQ scores at the end of grade 12. That is, this is an instance of the use of the experimental IAQ measure as a criterion of creativity. Relatively complete data on over 1,000 students lead to the following conclusions to date. The best predictors of present and future creative achievement were measures of past achievement. Seventh grade academic skills seem to correlate significantly with the criteria of creativity under study. The combination of past achievements and academic skills had correlations ranging between .40 and .50 with IAQ scores. And the particular criterion area—for example, creativity in science versus creativity in art—and the sex of the student seemed to be the chief moderators of the relationships found in this study. The research should be of interest to researchers in creativity and developers of noncognitive ability measures.

#### *Models for Measurement and Learning.*

ERIC HOLMAN

Two lines of research were pursued in this study: (1) the application of conjoint measurement to asymptotic learning data, and (2) the investigation of sequential properties of two-choice learning. Among the findings: Although there are exceptions that warrant further investigation, the conjoint measurement axioms seem to be satisfied to a first approximation by most of the available learning data. Consequently, the asymptotic response probabilities observed in simple learning experiments can be ordered according to the

product of two functions in the case of classical conditioning, and three functions in the case of instrumental conditioning. As for the second line of investigation: When used to distinguish between learning models, most sequential statistics contain relatively few degrees of freedom. Moreover, and rather surprisingly, these sequential statistics tend to reflect the total distribution of response probabilities over all trials, rather than trial-by-trial changes in response probabilities. The research should be of interest to researchers in psychometrics and learning theory.

#### *Error Choice Study.*

THOMAS S. BARROWS

One of the most difficult problems in measuring attitudes is how to construct measures that will uncover what the person *really* feels or believes, rather than what he senses may be socially acceptable. This study is exploring the use of the error choice technique in measuring attitudes toward Negroes. The approach uses an item format similar to that of standard, factual multiple-choice questions; however, with the error choice technique there is no correct answer *per se*. Each choice for any multiple-choice question is designed to reflect varying degrees of attitude. The study should be of interest to psychologists studying attitude measurement, and to psychometricians studying characteristics of various item types. It should be particularly relevant to the interests of those concerned with developing more dependable ways of measuring attitudes toward Negroes both in educational settings and more generally.

#### *Cross-Sectional vs. Longitudinal Results.*

THOMAS L. HILTON, CATHLEEN PATRICK

The ETS Growth Study has a data bank of the results from repeated testings of 32,000 students in 17 U.S. cities from 1961 to 1967. A difficulty in making use of this rare collection of information has been the problem of relating longitudinal data of this nature to cross-sectional study results—so that one may safely generalize. Longitudinal data, for example, are particularly val-

uable because they are compiled on the same individuals over an extended period of time. Such data are gathered, however, from a limited subsample of the population sampled by the cross-sectional approach. This study sought to determine relationships between the two research methods by comparing mean achievement test scores for matched longitudinal, unmatched longitudinal, and cross-sectional data. It found that matched longitudinal data yielded significantly higher means than either unmatched longitudinal or cross-sectional data. It also found that a school's drop-out rate was highly related to the fifth grade discrepancy in mean scores between cross-sectional and matched longitudinal data ( $r = .85$ ). The study should be of interest to researchers studying individual development and evaluating school effects, and to school administrators and policy makers attempting to interpret test results of groups of students.

#### *Review of Literature on Self-Fulfilling Prophecy.*

WILLIAM E. COFFMAN, MARY E. PARRY

Sociologist Robert Merton coined the phrase "self-fulfilling prophecy" for a phenomenon that has been noted in many contexts—the strange fact that quite often things happen according to one's expectations. The self-fulfilling prophecy has been discerned at work in bank failures, bowling games, warfare, and, within education, in low achievement by students in apparent response to low expectations of teachers. Of major concern to researchers are the possible effects of expectation in distorting research findings. This study involved a search of the literature, mainly in the fields of psychology and education, seeking to identify problems involving experimenter bias in research, as well as ways to control for, or prevent, such bias. The study identifies numerous instances of biasing expectations on the part of the research investigator and his subjects (the Hawthorne effect, for example) as well as numerous controls. Since the problem occurs in research of almost every type, this study should be of interest to researchers in all areas.



*Decision Theory Approach to Multi-Stage Testing.*

DONALD A. ROCK, ROBERT L. LINN,  
T. ANNE CLEARY

This study has somewhat the same aim as the study of programmed tests—to reduce testing time by applying sequential branching techniques. However, there is an important difference. Rather than branching to fit the individual with each question—as, in a sense, the programmed test approach does—in this case, the individual is presented with a choice of many short tests, each of which measures a different aspect or level of his abilities. The choice of the particular set of tests that best fits him is determined by a sequence of questions—his answers determine which short tests he should take, and which he shouldn't, as well as how many short tests of a particular type he should take before his ability in any one area is accurately assessed. Progress has been made in developing classification methods and computer techniques. The research should be of interest to test development specialists and computer scientists. RB-68-11.

*An Investigation of Student Test-Taking Behavior.*

DEAN W. SEIBEL, SCARVIA B. ANDERSON,  
ARLEEN S. BARRON

An early step in test development is pretesting to see whether the test measures what it is supposed to, and to find out how it may be improved. Statistical data (such as p-values and discrimination indices) are generally used to evaluate pretesting results. However, a persistent question is whether overreliance on statistical measures may blind test developers to other important clues to test relevance and usefulness. This study sought to find out: (1) whether it is possible to increase the value of pretesting operations by including observational as well as statistical data, (2) whether children of differing social classes show observable differences in test-taking behavior, and (3) whether other classroom characteristics are test relevant. The method is of particular interest in that a "natural" observer (the child's regular teacher) and an external observer (an

ETS researcher) observed how students of different social classes (high or low SES) responded to various test items. Forty third-grade, 82 seventh-grade, and 92 eleventh-grade classes, and more than 2,900 different items were involved. The data are being analyzed. When finished, the study should be of considerable interest to test developers and to all those concerned with evaluation that will capture the interest—as well as more fairly evaluate—children of all backgrounds.

*Cognitive and Affective Dimensions of Differential Achievement on Open- and Closed-Book Achievement Tests.*

GARY L. MARCO

Many questions confront a teacher when he weighs the considerations for giving an open-book versus a closed-book exam. This study examined these three: Will taking an open-book test result in higher achievement scores? Will it result in lower test anxiety? And what cognitive and personality variables are related to open-book achievement testing? The subjects were 164 college seniors. It was found that the open-book exam increased their achievement on knowledge items but not on application items. The open-book exam also seemed to decrease anxiety slightly. As for test-relevant abilities, the same types seemed to be required for both open- and closed-book exams. The research should be of interest to teachers. It suggests they can use open-book exams to help reduce anxiety without markedly affecting achievement.

*Goal-Oriented Teaching Exercises.*

DONALD M. MEDLEY

Needs for better methods of training teachers and better means of measuring teachers' on-the-job effectiveness have long been recognized. The aim of this project is to meet both needs, as well as to aid research. The approach is to obtain for analysis a standardized sample of a teacher's classroom behavior by means of the Goal-Oriented Teaching Exercise. This is a brief teaching unit, defined in terms of specified out-

comes that can be measured by a test. A four-day unit on air pollution was developed. The unit was used by teachers in eight junior high school general-science classes with both the teachers' and the students' behavior being recorded by audio and video tape recordings. These tapes were analyzed to see if ways of relating specific types of teacher behavior to definite evidence of pupil learning could be found. Results strongly suggest that even when they teach the same unit, different teachers aim for and achieve different kinds of pupil learning. The project was carried out in collaboration with R. A. Hill of Temple University. Its results should be of interest to teachers, administrators, and educational researchers.

*SCAT-STEP Factor Structure.*

KARL JORESKOG, COSTA W. BERGLUND

The fact that some of our abilities remain fairly constant while others change as we grow older poses problems for test developers. They must, for instance, gain an understanding of what may be happening to the underlying factor structure of any series of tests that is to measure apparent growth in skills over a sequence of years. The purpose of this study is to investigate stability and change over time of the factor structure of two ETS tests designed for use from grades 5 through 11, the School and College Ability Tests and the Sequential Tests of Educational Progress. Data from the total Growth Study sample are being analyzed. The approach involves fitting of alternative factorial models to test scores obtained at grades 5, 7, 9, and 11. Being taken into consideration are general factors across all tests, factors specific to individual tests across grades, and factors specific to particular grades. When completed, the research should be of interest to test developers and to primary and secondary school administrators and teachers.

*Differential Weighting of Item Distractors in Data Sufficiency Items.*

THOMAS F. DONLON, GEORGE H. DUNTEMAN

An important practical consideration



TABLE 1

## Behaviors Recorded In The First Word

| Symbol  | Name                               | Brief Description                                                                                |
|---------|------------------------------------|--------------------------------------------------------------------------------------------------|
| 1. PNS  | Pupil Utterance<br>Non-Substantive | Pupil makes a statement or asks a question not related to substance                              |
| 2. PQU  | Pupil Question<br>Substantive      | Pupil asks for substantive information                                                           |
| 3. PST  | Pupil Statement<br>Substantive     | Pupil offers substantive information                                                             |
| 4. PRS  | Pupil Response                     | Pupil responds directly to another pupil or indirectly to teacher                                |
| 5. PBST | Problem-Structuring<br>Statement   | Teacher raises a substantive question or sets a problem (without indicating who is to answer it) |
| 6. CVG  | Convergent Question                | Teacher asks pupil a question, which calls for one right answer                                  |
| 7. EL1  | Elaborating 1 Question             | Teacher directs question to the same pupil who answered the question preceding it                |
| 8. EL2  | Elaborating 2 Question             | Teacher directs question to pupil whose answer depends on the preceding one                      |
| 9. DVG  | Divergent Question                 | Teacher asks pupil a question to which more than one answer may be acceptable or correct         |

Teacher behavior is being measured by Donald Medley. The tables on this and the next page explain the symbols used in the rating form for classroom observers, known as OScAR (Observation Schedule and Record Form). The potential interest of these brief lists is that researchers find these observable bits of behavior particularly useful in assembling descriptions that appear to be relevant to selecting, training, and assessing the performance of teachers. (See Medley, *Goal-Oriented Teaching Exercises*.)

TABLE 2

## Behaviors Recorded In The Second Word

| Symbol   | Name                                         | Brief Description                                                                              |
|----------|----------------------------------------------|------------------------------------------------------------------------------------------------|
| 1. NOEV  | No Evaluation                                | Teacher does not reply to pupil utterance                                                      |
| 2. CNSUP | Considering-Supporting                       | Teacher utterance with positive affect                                                         |
| 3. INFAP | Informing-Approving                          | Teacher gives information or positive feedback                                                 |
| 4. DSCAC | Describing-Accepting                         | Teacher accepts pupil response or makes statement not otherwise classifiable                   |
| 5. DIREJ | Directing-Rejecting                          | Teacher commands pupil to do something or gives negative feedback                              |
| 6. RBCRT | Rebuking-Criticizing                         | Teacher utterance with negative affect                                                         |
| 7. DST   | Desisting                                    | Teacher commands pupil to stop doing something or refuses permission                           |
| 8. PRNS  | Procedural, Neutral-Non-Substantive Question | Teacher asks question not otherwise classifiable; teacher neither refuses nor gives permission |
| 9. PR+   | Procedural, Positive                         | Teacher offers pupil choice of action or gives permission                                      |

in test development is test efficiency, in the sense of gaining useful information per unit of testing time. That is, if everybody gets roughly the same score on an item it will be of little use, so the objective is to find questions that some can answer and some cannot. But then what about the case of the test-taker who knows some of the answer, but not all of it? Shouldn't he gain some credit for this knowledge, rather than be simply ruled right or wrong? This study sought to find out whether a test of this type might be developed by using a scoring scheme that assigned different weightings to wrong answer options. That is, a partially correct answer would get some bit of credit more than a wholly wrong answer. It was found that using the approach did not yield appreciable gains in useful variance; however, the method for assigning weights was partially successful, and the study revealed ways of possibly improving the method of weighting.

The research should be of interest to psychometricians and test developers. CEEB.

#### *Bayesian Methods of Item Allocation.*

ROGER OWEN

The purpose of this research is to use Bayesian methods to determine the most useful sequence or arrangement of items in a test. Two research bulletins describe the results. "A Bayesian Approach to Tailored Testing" considers the sequential design and analysis of a test consisting of dichotomously scored items. "The Optimum Design, Using Prior Information, of an Examination for Inference about Items with a Normal Additive Model" investigates the extent to which an examinee's expected observed score (or true score) on an item depends on his ability and on the easiness of the item. The results will be of interest to test developers and psychometricians. RB-69-92.

#### *Estimation of Test Weights by the Subjective Utility Function.*

DONALD SCHEWARTZ

There is a need in applied psychological measurement for some simple method

to determine optimal test weights for use in situations where validity studies are unfeasible for economic or other reasons. This study attempted to see if such weights could be estimated using subjective utility scale values in a judgment situation. The subjects are asked to judge whether or not people who have been selected by a particular test battery are "successful." Since subjective utility scale values can be determined by a simple and inexpensive questionnaire, this method might provide an economical substitute for costly predictive validity studies. A group of candidates for promotion to Senior Civil Engineering Assistant were tested and evaluated. It was concluded that (1) a subjective utility function does exist over a finite series of criterion elements in a personnel selection problem and can be measured by scaling methods, (2) this utility function is a good approximation of the optimal test (beta) weights in an idealized setting, and (3) the practical application of this method of weighting tests requires the use of tests with high construct validity.

## *Computer Studies*

#### *Exploratory Project on Adaptive Programming.*

ERNEST J. ANASTASIO

In recent years, complex computer systems have been developed for a variety of educational uses. A number of weaknesses in these systems have become apparent, however, which seem solvable only through more research in computer simulation of human processes. Thus, a major goal of this project was to construct systems, in the form of computer programs, that would "behave rationally" by simulating the outcome or product of the perceptual and reasoning processes used by humans in problem-solving situations. Efforts to date have yielded three functioning programs, each with varying degrees of true adaptive capability. One is a game-playing program that

competes against a human—the researcher attempts to outdo the computer playing the complex game of Go, for example. Another is a hypothesis testing program that demonstrates how the computer can be used cooperatively, rather than competitively, in interactive problem solving. A third program, like the first, is also competitive, but in this instance the computer's "playing ability" can be adjusted to any one of nine levels to provide a better match for the skills of its human opponent. So far, aside from the development of programs, it has become apparent that a central, and possibly key, requirement of adaptive programming, is the translation of human thought processes into the algorithmic and heuristic forms readily compatible with computer processing. The research should be of interest to computer-

assisted instruction specialists, curriculum designers, and computer scientists.

#### *Computer Generation of Test Items.*

ERNEST J. ANASTASIO

Considerable progress has been made by computer technicians and test specialists in developing systems that make it possible to turn over to the computer much of the task of assembling test items (or questions) into tests. However, the items must have been previously classified and pretested. The general purpose of this study was to determine whether computer assistance was feasible in the preparation of test items. A more specific objective was to identify the properties of "good" test items for measuring verbal ability, and to develop rules for coding words



PAGE 1

```

      DEFINE( 'REPLACE(KEY,INSERT)', 'START')
BEGIN LIST = TRIM(SYSPIT)
      LIST 'E*N*D' /S(END)
A      LIST ' ' = ' ' /S(A)
B      LIST ' ' = ' ' /S(B)
C      LIST '*' = ' ' /S(C)
D      LIST 'QUOTE' = ' ' /S(D)
LOOPA LIST *WORD* ' ' *LIST* /F(BEGIN)
      MODE('ANCHOR')
      LENGTH = SIZE(WORD)
      .LT( LENGTH , '5' ) /S(LOOPA)
      SAVE = WORD
*RULES FOR BEGINNING OF WORD ONLY
      POINTER = '0'
      REPLACE('PRE','PER')
      REPLACE('PER','PRE')
      REPLACE('ADJ','AJ')
      REPLACE('AJ','ADJ')
      REPLACE('ACQ','AQ')
      REPLACE('AQ','ACQ')
*RULES FOR EVERYWHERE EXCEPT LEFT AND RIGHT TERMINALS
LOOPB POINTER = POINTER + '1'
      WORD */'1'* *WORD*
      REPLACE('IE','EI')
      REPLACE('EI','IE')
      REPLACE('MB','MM')
      REPLACE('SH','S')
      REPLACE('CH','C')
      REPLACE('SC','S')
      REPLACE('SC','C')
      REPLACE('EA','EE')
      REPLACE('EE','EA')
      REPLACE('OR','ER')
      REPLACE('ER','CR')
      REPLACE('XH','H')
      REPLACE('PH','F')
      REPLACE('WR','R')
      REPLACE('GHT','HGT')
      REPLACE('ANC','ENC')
      REPLACE('ENC','ANC')
      REPLACE('AGE','EGE')
      REPLACE('EGE','AGE')
      REPLACE('ANT','ENT')
      REPLACE('ENT','ANT')
      REPLACE('ABL','EBL')
      REPLACE('EBL','ABL')
      REPLACE('ABL','IBL')
      REPLACE('IBL','ABL')
      REPLACE('EBL','IBL')
      REPLACE('IBL','EBL')
      REPLACE('ION','IN')
      REPLACE('OUL','OL')
      REPLACE('OUL','UL')
      REPLACE('PHY','PY')
      REPLACE('EUR','URE')
      REPLACE('CHO','CC')
      REPLACE('PRE','PER')

```

This is a portion of a program written in the SNOBOL programming language that was used to generate spelling items in a computer-assisted item-writing project. The program systematically *misspells* words in order to create distractor items. (See Anastasio, *Computer Generation of Test Items*.)

and sentences so they can be manipulated by a computer. After a systematic analysis of the nature of discrete verbal items—spelling and sentence completion items, for example—the project was able to show that the computer can carry out some of the less creative steps in item writing. It also concluded that systematic exploration of the item-writing process could make possible the computer generation of test items with prespecified properties—that is, items that would require only minor editing before preliminary field testing. Though the study uses rather well-known computer techniques, this was the first application to this particular educational problem. The research should be of interest to test developers, computer-assisted instruction specialists, and computer scientists. TDM-68-1, TDM-69-1, *Journal of Educational Measurement*, Vol. 6, No. 2, Summer 1969.

#### *Computer-Assisted Testing.*

ROBERT L. LINN, DONALD ROCK

Using computer consoles for testing examinees could make possible a number of improvements in testing procedures. For example, it makes possible the tailoring of items and their sequencing to the individual examinee. It also permits the use of new types of questions that require an interactive exchange between the computer and the examinee. This study is evaluating new testing procedures of this type, which would make use of the flexibilities offered by computer-assisted testing devices. Though the research is still in the data collection phase, the results could be useful to a number of ETS testing programs. The research is also being related to research in mental test theory and in the areas of human problem solving and computer-assisted instruction. It could lead to more efficient measurement, and provide a means for measuring dimensions not currently measured by paper and pencil tests—for example, problem solving strategies that might be important to the choice of instructional methods. A report appeared in *Educational and Psychological Measurement*, 1969, 24, 134-145. CEEB. (T. A. Cleary, co-author.)

#### *A Computer Model of Everyday Human Thinking.*

WILLIAM A. LEAF

A use of the computer that has greatly intrigued researchers over the past decade has been the way it may be used to simulate human thinking processes. Much may be learned about human thinking if one can translate the various operations involved into a set of instructions to the computer, and then have the computer solve problems in roughly the same sequences and with roughly the same results. This study used recent research and theorizing to outline a model of the processes people use in normal thinking. Specifically, the model was that which seemed to be used by a person in maintaining a valid mental picture of the outside world and its various relationships to him. The model was written as a computer program (in a list-processing extension of PL-I) to force specifying the steps that are assumed or implicit and thus operationally vague in most formulations. After being coded, the model will be tested for the degree to which it can duplicate a real-life, problem-solving situation. When completed, the research should be of interest to those interested in computer simulation of human thinking processes, or "artificial intelligence," as it is known.

#### *Multivariate Statistical Analysis.*

HARRY H. HARMAN

One reason for the popularity of the computer among researchers is that it makes possible carrying out within minutes statistical analyses of data that would otherwise take weeks or months. However, in order to gain this dramatic advantage in time, it is necessary to plan for it in one's research and to have a range of computer programs that offers some flexibility of choice. This project has concentrated on developing such a range of computer programs within the very broad area of multivariate statistical analysis. Particular emphasis has been placed on the development and adaptation of programs for studies that are still in relatively early, unstructured, probing

stages. A manual of scientific programs has been prepared to communicate the availability of these programs. It now contains about a dozen computer programs in the areas of multivariate statistical analysis and operations research, and additional programs are being included in the manual as they become available. The project should be of interest to all researchers making use of multivariate statistical analysis.

#### *Computer-Assisted Testing Symposium.*

HARRY H. HARMAN, CARL E. HELM,  
DAVID E. LOYE

In the eruption of acronyms that characterize the computer revolution, two names have emerged from the educational setting about which some confusion exists: CAI, or computer-assisted instruction, and CAT, or computer-assisted testing. Those working on instruction find they are also involved in testing, and those working in testing find they are also involved with instruction. To explore the possibilities of a productive interaction between the two interests, and also to help shape long-term goals in this area for ETS, a research symposium was convened at ETS. For two days computer scientists from organizations throughout the country probed issues in the field and the proceedings were transcribed, edited, and have been published in *Computer-Assisted Testing*, available through the ETS Office of Computation Sciences. Participants included Ernest Anastasio, A. G. Bayroff, Gordon Becker, Richard Cabot, Garlie Forehand, Harry Levitt, Robert Linn, H. William Morrison, Allen Newell, Warren T. Norman, H. G. Osburn, Donald Rock, Robert Ruderman, Emir Shuford, Keith Smith, and John Vinshonhaler. Discussion leaders and symposium organizers were Harry Harman and Carl Helm.

#### *Model for ETS Shipping.*

LOUIS R. LAVINE

A fascinating, relatively new pursuit within the educational community is the use of simulation models to cut costs and improve the efficiency of various operations. In essence, the total operation of some process is



analyzed from start to finish, and this operation is then translated into programs for the computer. This programmed, computerized "replica" of the actual real life operation is known as a simulation model. A sizable management problem for ETS is scheduling the shipping of its tests throughout the United States and throughout much of the world. If optimal shipping strategies can be developed for ETS testing programs, overshipment of test materials can be avoided, and other important economies can be effected. This study designed and programmed a simulation model that will permit ETS program directors, analysts, and others to evaluate alternative strategies for shipping test materials. Strategies for shipping test materials that might be far too costly to compare in real life can be quickly and cheaply compared in the model.

#### *Scheduling Techniques.*

LOUIS R. LAVINE

The scheduling of workloads in large organizations has gone beyond being a fine art practiced by adept managers. It is becoming a fine science practiced according to instructions from the computer. This study applied standard Program Evaluation and Review Technique (PERT) procedures to activities ranging from the Longitudinal Study of Disadvantaged Children to conversion from second-generation to third-generation computers. As an independent effort, the study also developed a work-

load scheduling program that utilized an algorithm for smoothing the overall workload for an individual (or an organization) by introducing peaks and valleys into the workload schedules for independent activities. The purpose of this effort was to develop techniques that may assist project directors and others in managing complex internal projects. It should be of interest to ETS project directors and division directors, and to supervisory and staff personnel in the operational services departments. RM-69-14.

#### *Multivariate Change Pilot Study.*

ALBERT A. BEATON, THOMAS L. HILTON,  
CATHLEEN PATRICK

Ever since the computer emerged on the scene, the hope has persisted within educational research that useful theoretical models of very complex educational processes might be constructed. Should the approach prove economically and technically feasible, one of many advantages would be that research results, which tend now to be lost in reports, would be accumulated. This study is exploring an approach of this type in making use of Growth Study data. Using measures and biographical information relating to science achievement, a computer program for a method of longitudinal multivariate analysis has been developed and is being tested on a number of theories of growth. Preliminary results indicate this may be an effective way of studying complex growth processes. The

research should be of interest to educational theorists and to researchers concerned with the application of the computer to education.

#### *A List Processing Extension to PL-I.*

WILLIAM A. LEAF

A relatively new computer language known as PL-I is described by IBM as an extremely sophisticated language combining algebraic power with the ability to deal with data in a wide variety of forms. Eventually, PL-I is intended to become the standard language for programmers whose applications currently require many different languages. However, PL-I has a weakness: its poor facility for dealing with alphanumeric or numeric data stored in variable-length lists. The purpose of this project was to write routines that will give programmers easy methods of generating and manipulating such lists while retaining the full power of PL-I. Many of these routines have now been worked out. Their features include the dynamic allocation of space to handle increasing storage requirements of increasing-size lists, simple and informative input/output of list structures, and convenient manipulation of data on the lists. Also available is a subroutine called SNOSCAN, which duplicates the basic SNOBOL text processing facilities for PL-I character strings. A manual describing features and uses of PL-I/SLIP and SNOSCAN is being prepared. The research should be of interest to programmers and computer scientists.



## IV

---

# PUBLICATIONS AND SPEECHES OF STAFF MEMBERS

July 1, 1968 — June 30, 1969

ALKER, HENRY A.

Rationality and achievement: A comparison of the Atkinson-McClelland and Kogan-Wallach formulations. *Journal of Personality*, 1969, 37, 207-224.

(With Julia Carlson & Margaret Hermann) Multiple-choice questions and student characteristics. *Journal of Educational Psychology*, 1969, 60, 231-243.

(With Nathan Kogan) Effects of norm-oriented group discussion on individual verbal risk taking and conservatism. *Human Relations*, 1968, 21, 393-405.

ANASTASIO, ERNEST J.

(With S. H. Evans) Misuse of analysis of covariance when treatment effect and covariate are confounded. *Psychological Bulletin*, 1968, 69, 225-234.

(With John J. Fremer, Jr.) Computer-assisted item writing—I (spelling items). *Journal of Educational Measurement*, 1969, 6, 69-74.

(With D. M. Marcotte & John J. Fremer, Jr.) Computer-assisted item writing—II (sentence completion items). Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles, California, February 1969.

ANDERSON, SCARVIA B.

The making of a pupil: Changing children into school children. Speech presented at the University of Chicago, July 1968.

The ETS-OEO longitudinal study of disadvantaged children. Speech presented at the National Council on Measurement in Education, November 1968.

BARROWS, THOMAS S.

Bluffton interterm: Evaluation. Project Report 69-7. Princeton, New Jersey: Educational Testing Service, 1969.

BEATON, ALBERT E.

Scaling criterion of questionnaire items. *Socio-Economic Planning Science* (Vol. 2), Great Britain, 1969.

Algorithms for data maintenance and computation of analysis of variance. *Statistical Computation*. New York: Academic Press, 1969.

Criteria for evaluation. Research Memorandum 69-9. Princeton, New Jersey: Educational Testing Service, April 1969.

Assessment programs and evaluation techniques. Paper presented at the Operations Analysis Workshop of the Association for Educational Data Systems. Philadelphia, Pennsylvania, November 1968.

The computer techniques used in the Educational Opportunities Survey. Research Memorandum 68-16. Princeton, New Jersey: Educational Testing Service, July 1968.

Some considerations of technical problems in the Educational Opportunities Survey. Research Memorandum 68-17. Princeton, New Jersey: Educational Testing Service, July 1968.

(With Norman Frederiksen, Ollie A. Jensen, & Bruce Bloxom) Organizational climates and administrative performance. Research Bulletin 68-41. Princeton, New Jersey: Educational Testing Service, 1968.

(With G. W. Mayeske, W. M. Cohen, & T. Okada) The role of the family in the development of student achievement and motivation. Technical Paper Number 2, June 1969, Department of Health, Education, and Welfare, U.S. Office of Education.

BOLDT, ROBERT F.

Concurrent validity of the PAA and SAT for bilingual Dade County high school volunteers. Research and Development Report 68-9, No. 3. Statistical Report 69-31. Princeton, New Jersey: Educational Testing Service, April 1969.

Discrepant predictor study: Admission Test for Graduate Study in Business. Princeton, New Jersey: Educational Testing Service, June 1969.

BUSSIS, ANNE M.

(With Jean H. Orost, M. Papenek, & Masako N. Tanaka) Teaching classification concepts to disadvantaged children. Paper presented at the meeting of the American Educational Research Association, Los Angeles, California, February 1969.

(With Diran Dermen) Measurement of analytic functioning. Disadvantaged children and their first school experiences. Progress Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.

(With Virginia C. Shipman) Further considerations concerning the measurement of verbal behavior. Progress Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.

(With Virginia C. Shipman) The impact of the family. Progress Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.

BROWN, MICHAEL W.

A comparison of factor analytic techniques. *Psychometrika*, September 1968, 33, 267-334.

(With Walter Kristof) On the oblique rotation of a factor matrix to a specified pattern. *Psychometrika*, June 1969, 34, 237-248.

CAHEN, LEONARD S.

The role of large-scale projects in curriculum evaluation with examples from the National Longitudinal Study of Mathematical Abilities. In Anna Dragositz (Ed.) *Curriculum innovations and evaluation*, Proceedings of the Association for Supervision and Curriculum Development Pre-Conference Seminar. Princeton, New Jersey: Educational Testing Service, March 8-9, 1968.

Introduction to *Evaluation activities of curriculum projects* Vol. 2, *AERA Monograph Series on Curriculum Evaluation*, By H. Grobman, Rand-McNally, 1968. (Also edited by L. S. Cahen)

Common problems in evaluating programs, materials, and curricula. Research Memorandum 68-21. Princeton, New Jersey: Educational Testing Service, 1968. Republished in a modified version under the title, Comparative problems in product and curriculum innovation. *Educational Product Report* (Special issue on evaluation: Theory and practice), February 1969, 2.

Contemporary approaches to evaluation in mathematics education. Paper presented at a seminar on Mathematics Education, School of Education, University of Georgia, Athens, Georgia, February 1969.

CAMPBELL, JOEL T.

(With Leon H. Belcher) An exploratory study of word associations of negro college students. *Psychological Reports*, 1968, 23, 119-134.

CARLSON, JULIA

(With Clive Davis) A cross-cultural study of the Müller-Lyer Illusion as a function of attentional factors. Research Bulletin 68-53. Princeton, New Jersey: Educational Testing Service.

(With Clive Davis) A cross-cultural study of risk taking. Paper presented at the annual convention of the African Studies Association, Los Angeles, California, October 16-19, 1968.

(With Nathan Kogan) Difficulty of problems attempted under conditions of competition and group consensus. *Journal of Educational Psychology*, 1969, 60, 158-167.

(With Henry Alker and Margaret Hermann) Multiple-choice questions and student characteristics. *Journal of Educational Psychology*, 1969, 60, 231-243.

CARROLL, JOHN B.

Contrastive analysis and interference theory. *Georgetown University Monograph Series on Languages and Linguistics*, 1968, 21 (*Research in Education*, 1969, 4(2), 19. Doc. No. ED 022158).

The psychology of language testing. In Alan Davies (Ed.), *Language testing symposium: A psycholinguistic approach*. London: Oxford University Press, 1968.

Psychological aspects of programmed learning in foreign languages. In Theodore H. Mueller (Ed.), *Proceedings of the Seminar on Programmed Learning*. University of Kentucky 20th Foreign Language Conference, Lexington, Kentucky, April 29, 1967. New York: Appleton-Century-Crofts, 1968.

Learning from verbal discourse in educational media: Some research studies. In Institute for Development of Educational Activities (Comp. and Ed.) *Project ARISTOTLE Symposium*. Washington: National Security Industrial Association, 1969.

Review of Holmes, J. A., and Singer, H. *Speed and power of reading in high school*. U. S. Office of Educ. Coop. Res. Program Monogr. No. 14 (*Research in the Teaching of English*, 1968, 2(2), 172-184).

European seminar on learning and the educational process. *International Newsletter Issue IV*, Princeton, New Jersey: Educational Testing Service, April 1968.

*Final Report on the European Seminar on Learning and the Educational Process* at Skepparholmen, Saltsjö-Boo, Sweden, July 29-August 23, 1968. Hamburg: UNESCO Institute for Education, November 1968.

(With H. Leslie Cramer) *The intelligibility of time-compressed speech, Final report*. Cambridge, Mass.: Laboratory for Research in Instruction, July 1968 (EDRS, Doc. No. ED 021 851).

Reflections on language and thought. (Discussion of Professor Jenkins' paper.) In James F. Voss, *Approaches to thought*. Columbus, Ohio: Charles E. Merrill Publishing Co., 1969.

Language and psychology. In Archibald A. Hill, (Ed.), *Linguistics today*. New York: Basic Books, Inc., 1969.

The European Seminar on learning and the educational process, Skepparholmen, Sweden, July 29-August 23, 1968. *Social Science Research Council Items*, 1969, 23(1), 6-10.

Phillip Justin Rulon (1900-1968). *Psychometrika*, 1969, 34, 1-3.

Psychological and educational research into second language testing to young children. In H. H. Stern (Ed.), *Languages and the young school child*. London: Oxford University Press, 1969.

Guide for the collection of data pertaining to the study of foreign or second languages by younger children. In H. H. Stern (Ed.), *Languages and the young school child*. London: Oxford University Press, 1969.

Psycholinguistics and the elementary language arts. In James Walden (Ed.), *Oral language and reading: Papers collected from the 1967 Spring Institutes on the Elementary Language Arts*. Champaign, Illinois: National Council of Teachers of English, 1969.

Reply to review, by Professor G. Herdan, of H. Kucera and W. N. Francis, *Computational analysis of present-day American English*. *General Linguistics*, 1968, 8(2), 156-157.

*From comprehension to inference*. Research Memorandum 69-11. Princeton, New Jersey: Educational Testing Service, May 1969.

The educational researcher: Computer scientist or Renaissance man? In Lorne H. Woollatt, (Ed.), *The emerging research worker and his needs: Report of the seventh Convocation on Educational Research, November 13-15, 1966*. New York State Education Department and the Educational Research Association of New York State, 1969.

A model of school learning. Speech given at the School of Education, University of Gothenburg, Sweden, August 26, 1968.

Studies of foreign language learning. Lecture given at the English Language Institute, University of Gothenburg, Sweden, August 24, 1968.

A rationale for the lognormal distribution of word-frequencies. Speech given at the Department of Psychology, State University of New York at Buffalo, December 3, 1968.

Discussant, Symposium: Confidence on achievement tests—Theory and application. Speech given at the American Educational Research Association Convention, Los Angeles, California, February 6, 1969.



- From comprehension to inference. Speech given at the Claremont Reading Conference, Claremont Graduate School, Claremont, California, February 7, 1969.
- A rationale for the lognormal form of word-frequency distributions. Speech given at the Eastern Psychological Association Convention, Philadelphia, Pennsylvania, April 11, 1969.
- The lognormal model of word-frequency distribution and its rationale. Speech given at the Institute of Experimental Psychology, Oxford University, England, May 13, 1969; University of Edinburgh, Scotland, April 14, 1968; Institute for Studies in Education, Toronto, Canada, April 14, 1969.
- A model of school learning and its educational implications. Speech given at the Department of Psychology, Louisiana State University, New Orleans, Louisiana, May 19, 1969.
- CASE, C. MARSTON  
The application of *PERT* to large-scale education research and evaluation studies. *Educational Technology*, October, 1969. (Speech presented at AERA, Los Angeles, California, February 1969).
- CASSERLY, PATRICIA L.  
What college students say about advanced placement I. *College Board Review*, Fall 1968-69.  
(With William E. Coffman) What college students say about advanced placement II. *College Board Review*, Winter 1968-69.  
Speech, Advanced Placement History Conference, Seattle, Washington, June 22, 1969.
- CENTRA, JOHN A.  
Instrumentation issues in studies of institutional environments. Speech presented at the Association for Institutional Research Conference, 1968.  
(With Rodney Hartnett) Attitudes and secondary school backgrounds of Catholics entering college. *Sociology of Education*, Spring, 1969, 42 (2).  
(With Robert L. Linn and Mary Ellen Parry) Academic growth in predominantly negro and predominantly white colleges. *American Educational Research Journal*, January 1970. Research Bulletin 69-39. Princeton, New Jersey: Educational Testing Service, May 1969.  
College freshman attitudes toward cheating. *Personnel & Guidance Journal*, January 1970. Research Bulletin 69-24. Princeton, New Jersey: Educational Testing Service, April 1969.  
(With Donald A. Rock and Robert L. Linn) The identification and evaluation of college effects on student achievement. *American Educational Research Journal*, January 1970. Research Bulletin 69-27. Princeton, New Jersey: Educational Testing Service, April 1969.  
(With Robert L. Linn) On the interpretation of student perceptions of their college environment. *Measurement and Evaluation in Guidance*, Summer 1970. Research Bulletin 68-34. Princeton, New Jersey: Educational Testing Service, August 1968.
- The development of the questionnaire on Student and College Characteristics. Research Memorandum 68-11. Princeton, New Jersey: Educational Testing Service, 1968.
- CHITTENDEN, EDWARD A.  
School achievement of first- and second-born siblings. *Child Development*, 1968, 39, 1223-1228.  
Some implications of Piaget's work for testing young children. Research Memorandum 69-7. Princeton, New Jersey: Educational Testing Service, 1969.  
Classroom structure and cognitive development: Evidence from recent studies. Paper presented at the annual Institute of Early Childhood Education Council and the Metropolitan Association for Childhood Education, New York, New York, March 1969.
- COFFMAN, WILLIAM E.  
Achievement tests. In R. L. Ebel (Ed.). *Encyclopedia of Educational Research*. Chicago: The MacMillan Co., 1969.  
Implications of national assessment for the school system. Paper presented at a symposium on National assessment: Progress and prospects, Joint Session of the AERA-NCME annual meeting, Los Angeles, California, February 6, 1969.  
Essay examinations. Research Memorandum 69-18. Princeton, New Jersey: Educational Testing Service, June 1969.  
Concepts of achievement and proficiency. Paper presented at the 1969 Invitational Conference on Testing Problems, New York, New York, November 1, 1969.
- CROOKS, LOIS A.  
(Ed.) Managing organizational stress. *Proceedings of the Executive Study Conference, November 29-30, 1967*. Princeton, New Jersey: Educational Testing Service, 1968.  
(Ed.) Managing organizational effectiveness. *Proceedings of Executive Study Conferences, May 14-15, 1968, November 13-14, 1968*. Princeton, New Jersey: Educational Testing Service, 1969.  
Issues in the development and validation of in-basket exercises for specific objectives. Research Memorandum 68-23. Princeton, New Jersey: Educational Testing Service, November 1968.  
Issues in the development and validation of in-basket exercises for specific objectives. Paper presented at meetings of the American Psychological Association, symposium on the In-basket exercise—Test or technique? San Francisco, California, September 1968.
- DIEDERICH, PAUL B.  
Progressive education should continue. *Today's Education*, 1969, 58 (3), 12-19.  
What does research in reading reveal about evaluation in reading? *The English Journal*, 1969, 58, 853-868.  
The transition from juvenile to adult books. *Educational Horizons*, 1968, 56, 153-158.  
The development of a national assessment program in English. *Research in the Teaching of English*, Spring 1969, 3, 5-14.

DYER, HENRY S.

Summary and evaluation of contributions to the yearbook: Race and equality in American education. *The Journal of Negro Education*, Summer 1968, 210-358.

Issues in testing. Speech presented at the Organization for Economic Cooperation and Development, Ford Foundation, New York, New York, January 15, 1969.

Research between the right and left. Speech presented at the RELCV Workshop, Durham, North Carolina, January 30, 1969.

An educational researcher's view of systems analysis. *Educational Technology*, 1969, 9 (9), 83-85.

The mission of the evaluator. *The Urban Review*, February 1969.

The agony and the promise of higher education as seen through ETS research. Paper presented at the American Association of Higher Education, Chicago, Illinois, March 3, 1969.

Systems analysis from the viewpoint of educational research. Paper presented at the 1969 Educational Leadership Conference, Manchester, New Hampshire, April 7, 1969.

EMMERICH, WALTER

Social role development. *Perspectives on Human Deprivation*. National Institute of Mental Health, 1968.

How are measurement strategies related to models of human development? Paper presented at the symposium on Untangling the tangled web of education, sponsored by the National Council on Measurement in Education, New York, New York, November 1, 1968. Research Memorandum 69-6. Princeton, New Jersey: Educational Testing Service, 1969.

Personality development and concepts of structure. *Child Development*, 1968, 39, 671-690.

Models of continuity and change in development. Paper presented at the biennial meeting of the Society for Research in Child Development, symposium on Behavioral continuity and change with development, Santa Monica, California, March 27, 1969.

EVANS, FRANK

(With Stephen F. Klein) An examination of the validity of nine experimental tests for predicting success in law school. *Educational and Psychological Measurement*, 1968, 28, 909-913.

(With Cathleen Patrick) Antecedents and patterns of growth of school dropouts. Paper presented at the 76th annual meeting of the American Psychological Association, San Francisco, California, August 30, 1968.

(With Stephen F. Klein & Donald A. Rock) The use of multiple moderators in academic prediction. *Journal of Educational Measurement*, 1968, 5(2).

FLAUGHER, RONALD L.

A fixed length optimal test battery for colleges characterized by diverse aptitude levels. *American Educational and Psychological Measurement*, 1968, 5, 659-674.

(With Richard S. Melton & Charles T. Myers) Item rearrangement under typical test conditions. *Educational and*

*Psychological Measurement*, 1968, 28, 813-824.

(With Donald A. Rock) A multiple moderator approach to the identification of over- and underachievers. Paper presented at the National Council on Measurement in Education, February 1969.

FREDERIKSEN, NORMAN

An experimental study of organizational climates and administrative performance. *Proceedings of the XVIth International Congress of Applied Psychology*, August 18-22, 1968. Amsterdam: Swets & Zeitlinger, 1969.

Administrative performance in relation to organizational climate. Paper presented at the annual meeting of the American Psychological Association Division 14 symposium on Measuring managerial effectiveness, San Francisco, California, September 1, 1968.

(With Stephen F. Klein) Anxiety and learning to formulate hypotheses. Paper presented at the American Educational Research Association, Los Angeles, California, February 8, 1969.

(With Bruce Bloxom) A three-mode component analysis of a person by situation by behavior data matrix. Paper presented at the Psychometric Conference honoring Paul Horst, Seattle, Washington, June 22, 1969.

FREEBERG, NORMAN E.

(With Donald A. Rock) Factorial invariance of biographical factors. *Multivariate Behavioral Research*, April 1969, 4, 195-210 (Paper also presented at the XVI International Congress of Applied Psychology, Amsterdam, August 1968).

Assessment of disadvantaged adolescents: A different approach to research and evaluation measures. Research Bulletin 69-41. Princeton, New Jersey: Educational Testing Service, May 1969.

Development of evaluation measures for the use with neighborhood Youth Corps enrollees. Progress Report 68-5. Princeton, New Jersey: Educational Testing Service, December 1968.

FREEDLE, ROY O.

Implications of a short-term memory method for a k-span learning model: Some data on children. Research Bulletin 69-4. Princeton, New Jersey: Educational Testing Service, 1969.

(With Marlys Craun) Observations with self-embedded sentences using written aids. Research Bulletin 69-22. Princeton, New Jersey: Educational Testing Service, 1969.

(With Terrance Keeney & N. Smith) Effects of phrase structure, mean depth, and grammaticality on children's imitation of sentences. Research Bulletin 69-79. Princeton, New Jersey: Educational Testing Service, 1969.

(With Terrance Keeney & N. Smith) Children's imitations of sentences evaluated by four response measures. *Proceedings of the 77th annual convention of the American Psychological Association*, 1969, 47-48.

(With Michael Lewis) On relating an infant's observation time of visual stimuli with choice theory analysis. Research Memorandum 69-1. Princeton, New Jersey: Educa-



- tional Testing Service, 1969.
- (With A. Zavala & E. A. Fleishman) Studies of component-total task relations: Order of component-total task practice and total task predictability. *Human Factors*, 1968, 10, 283-296.
- A modified k-span learning model based on a new short-term memory method. Paper presented at the Mathematical Psychology meetings, Stanford University, August 28-29, 1968.
- Asymmetric recognition matrices: A choice analysis as evidence for a hypothetical stimulus-sampling process. Paper presented at the spring meetings of the Psychometric Society, Educational Testing Service, Princeton, New Jersey, April 17-18, 1969.
- FREMER, JOHN J. JR.
- (With William E. Coffman) The College Board Scholastic Aptitude Test as a predictor of academic achievement in secondary schools in England. *Journal of Educational Measurement*, 1968, 5, 235-241.
- (With William E. Coffman) A look at the Scholastic Aptitude Test—1969. Speech presented at the workshop of Union County, New Jersey, Regional School District English teachers, February 1969.
- (With Ernest J. Anastasio and D. M. Marcotte) Computer assisted item writing—II (sentence completion items). Paper presented at the annual meeting of the National Council on Measurement in Education, Los Angeles, California, February 1969. Test Development Memorandum 69-1. Princeton, New Jersey: Educational Testing Service, 1969.
- (With Walter H. MacGinitie) The content of controlled continuous word associations. In J. Rosenstein, and W. H. MacGinitie (Eds.) *Verbal behavior of the deaf child: Studies of word meanings and associations*. New York: Teachers College Press, 1969.
- (With Ernest J. Anastasio) Computer assisted item writing—I (spelling items). *Journal of Educational Measurement*, 1969, 6, 69-74.
- FRENCH, JOHN W.
- Review of Raymond B. Cattell (Ed.). *Handbook of Multivariate Experimental Psychology*. *Psychometrika*, 1969, 34, 252-256.
- GOLDMAN, KARLA S.
- Attitudes & interests. *Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies*. Report in 2 volumes, December 1968, Contract OEO 4206, Office of Economic Opportunity. Progress Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.
- Taxonomy of children's interests. *Disadvantaged children and their first school experiences: Theoretical considerations and measurement strategies*. Report in 2 volumes, December 1968, Contract OEO 4206, Office of Economic Opportunity. Progress Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.
- GULLIKSEN, HAROLD
- Louis Leon Thurstone, experimental and mathematical psychologist. *American Psychologist*, 1968, 23(11), 786-802.
- Methods for determining equivalence of measures. *Psychological Bulletin*, 1968, 70(6), 534-544.
- Harold Gulliksen. a biography with bibliography. *Revue de Psychologie Appliquée*, 4th trim. 1968, 18(4), 241-244.
- Developments in mental testing, and their relation to learning. In W. Correll and Fritz Süllwold (Eds.), *Forschung und Erziehung: Untersuchen der Problemen der Pädagogik und Pädagogischen Psychologie*. Gesamtherstellung: Ludwig Auer, Donauwörth, 1968.
- HALE, GORDON A.
- Reversal shift performance in children as a function of age, overtraining, and dimension preference. Unpublished doctoral dissertation, University of Minnesota, 1969.
- The effect of dimension "dominance" on the overtraining reversal effect in children. Paper presented at the meetings of the Society for Research in Child Development, Santa Monica, California, 1969.
- (With L. K. Miller, & H. W. Stevenson) Learning and problem solving by retarded and normal subjects. *American Journal of Mental Deficiency*, 1968, 72, 681-690.
- (With G. M. Olson, L. K. Miller, & H. W. Stevenson) Long-term correlates of children's learning and problem solving behavior. *Journal of Educational Psychology*, 1968, 59, 227-232.
- (With H. W. Stevenson, R. E. Klein, & L. K. Miller) Interrelations and correlates in children's learning and problem solving. *Monographs of the Society for Research in Child Development*, 1968, 33, 7.
- (With H. W. Stevenson, R. E. Klein, & L. K. Miller) Solution of anagrams: A developmental study. *Child Development*, 1968, 39, 905-912.
- HALL, CHARLES E.
- A procedure for testing the homogeneity of dispersion matrices. Paper presented at a meeting of the Psychometric Society, April 1969.
- HARMAN, HARRY H.
- Factor analysis. Dean K. Whitla (Ed.), *Handbook of measurement and assessment in behavioral sciences*. Reading, Massachusetts: Addison-Wesley, 1968.
- (With Dean W. Seibel, Michael Rosenfeld, and Benjamin Shimberg) Evaluation of driver education and training. Progress Report 69-1. Princeton, New Jersey: Educational Testing Service, 1969.
- HARTNETT, RODNEY T.
- Difference in selected attitudes and college orientations between black students attending traditionally negro and traditionally white institutions. Research Bulletin 69-48. Princeton, New Jersey: Educational Testing Service, 1969.
- College and university trustees: Their backgrounds, roles, and educational attitudes*, Princeton, New Jersey: Educational Testing Service, 1969.
- (With Carol U. Libby) Agreement with views of Martin



- Luther King, Jr. before and after his assassination. Research Bulletin 68-35. Princeton, New Jersey: Educational Testing Service, 1968. (Eastern Psychological Association, Philadelphia, April 1969.)
- (With John A. Centra) Attitudes and secondary school background of Catholics entering college. *Sociology of Education*, 1969, 42(2).
- (With Richard Peterson) Religious preference as a factor in attitudinal and background differences among college freshmen. *Sociology of Education*, 1968, 41(2).
- HAVEN, ELIZABETH W.  
Curricular innovation and the subject-matter preparation of College Board test candidates. Speech presented at the American Educational Research Association, Los Angeles, California, February 5-8, 1969.
- HERMANN, MARGARET  
(With Henry A. Alker & Julia A. Carlson) Multiple-choice questions and student characteristics. *Journal of Educational Psychology*, 1969, 60, 231-243.
- (With J. A. Robinson & C. F. Hermann) Search under crisis in political gaming and simulation. In D. G. Pruitt & R. C. Snyder (Eds.), *Theory and research on the causes of war*. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.
- (With Nathan Kogan) Effects of representative status and decision style on cooperation in the prisoner's dilemma. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, Pennsylvania, April 1969.
- HILTON, THOMAS L.  
Ten basic questions in guidance and school learning. *Computer-based Vocational Guidance Systems*. Washington, D.C.: U. S. Government Printing Office, 1969.
- (With Cathleen Patrick) Cross-sectional versus longitudinal data: An empirical comparison of mean differences in academic growth. Research Bulletin 69-28. Princeton, New Jersey: Educational Testing Service, 1969.
- (With Michael Rosenfeld) Negro-white differences in adolescent educational growth. Research Bulletin 69-71. Princeton, New Jersey: Educational Testing Service, 1969.
- HOLMAN, ERIC W.  
General constraints on sequential properties of learning. Paper presented at the Mathematical Psychology meeting, Stanford, California, August 1968.
- IRONSIDE, RODERICK A.  
Demonstration reading classes. An ETS critical appraisal for the ESEA staff, Chicago Public Schools, July 1968.
- A test of literary appreciation for middle graders. Speech presented at the annual convention of the National Council of Teachers of English, November 1968.
- Measuring adult basic education skills: Emphasis on validity. Speech presented at the tenth national seminar on Adult Education Research, Toronto, Canada, February 1969.
- Who assesses reading progress: Tests, teachers, or students? Speech presented at the annual convention of the College Reading Association, Boston, Massachusetts, March 1969.
- Techniques in the diagnosis of reading difficulties. Speech presented at the annual meetings of the North Carolina Council, International Reading Association, April 1969.
- Law enforcement survey. Project report to the WGBH Educational Foundation, Boston, Massachusetts, July 1968.
- JACOBS, PAUL I.  
(With Mary Vandeventer) Progressive matrices: An experimental developmental, nonfactorial analysis. *Perceptual and Motor Skills*, 1968, 27, 759-766.
- (With Mary Vandeventer) The learning and transfer of double-classification skills by first-graders. Research Bulletin 68-65. Princeton, New Jersey: Educational Testing Service, 1968.
- (With Mary Vandeventer) Evaluating the teaching of intelligence. Research Bulletin 69-20. Princeton, New Jersey: Educational Testing Service, 1969.
- (With Mary Vandeventer) The learning and transfer of double-classification skills: A replication and extension. Research Bulletin 69-88. Princeton, New Jersey: Educational Testing Service, 1969.
- (With Mary Vandeventer) Information in wrong responses. *Psychological Reports*, 1970, 26, 311-315.
- JENSEN, OLLIE A.  
Constructing objective test items. In J. J. Donovan (Ed.), *Recruitment and selection in the Public Service*. Chicago, Illinois: Public Personnel Association, 1968.
- (With Norman Frederiksen and Albert E. Beaton) Organizational climates and administrative performance. Research Bulletin 68-41. Princeton, New Jersey: Educational Testing Service, 1968.
- JORESKEG, KARL G.  
Statistical models for congeneric test scores. *Proceedings of the American Psychological Association, 76th Annual Convention*, 1968, 213-214.
- (With D. N. Lawley) New methods in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology*, 1968, 21, 85-96.
- (With P. B. Warr and R. E. Lee) A note on the factorial nature of the F and D scales. *British Journal of Psychology*, 1969, 60, 119-123.
- A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 1969, 34, 183-202.
- Efficient estimation in image factor analysis. *Psychometrika*, 1969, 34, 51-75.
- KATZ, MARTIN R.  
Learning to make wise decisions. In Alice Scates (Ed.), *Computer-based Vocational Guidance Systems*. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1969.
- Counseling—Secondary school. In R. Ebel (Ed.), *Encyclopedia of Educational Research*. (4th ed.) New York: MacMillan, 1969.
- Theoretical foundations of guidance. *Review of Educational Research*, April 1969, 39, 127-140.

Are aptitude tests unfair to Negroes? In E.F. Borgetta (Ed.), *Social Psychology: Readings and Perspective*. Chicago: Rand-McNally, 1969.

Measuring interests. Tape recording produced for Sound Seminars, McGraw Hill, 1969.

Selecting an achievement test: Principles and procedures. Reprinted in N. Gronlund (Ed.), *Readings in Measurement and Evaluation*. New York: MacMillan, 1968.

*Tu: Oggi e Domani*. Firenze: Adattamento Italiano a Cura di Franca Magnani, 1968.

*Guida per L'Insegnante*. Firenze: Adattamento Italiano a Cura di Franca Magnani, 1969.

Rationale for a computer-assisted system of interactive guidance and information. Paper presented at the annual meeting of the College Entrance Examination Board, New York, New York, October 29, 1968.

KEENEY, TERRENCE J.

(With J. R. Jenkins & J. J. Jenkins) Identification of absolute and relational properties in the three-stimulus (ABB) configuration. *Child Development*, 1969, 40, 1223-1231.

Permutation transformations on phrase structures in letter sequences. *Journal of Experimental Psychology*, 1969, 82, 28-33.

Young children's imitation of nonsense syllable strings, word strings, and grammatical sentences. Paper presented at the meeting of the Society for Research in Child Development, Santa Monica, California, March 28, 1969.

The operation of permutation rules on hierarchical and non-hierarchical structures. Paper presented at the meeting of the Mid-Western Psychological Association, Chicago, Illinois, May 1969.

KOGAN, NATHAN

(With E. Pankove) Creativity and risk taking in elementary school children. *Journal of Personality*, 1968, 36, 420-439.

(With Julia Carlson) Difficulty of problems attempted under conditions of competition and group consensus. *Journal of Educational Psychology*, 1969, 60, 158-167.

(With W. Doise) Effects of anticipated delegate status on level of risk taking in small decision-making groups. *Acta Psychologica*, 1969, 29, 228-243.

(With Henry A. Alker) Effects of norm-oriented group discussion on individual verbal risk taking and conservatism. *Human Relations*, 1968, 21, 393-405.

(With Margaret G. Hermann) Effects of representative status and decision style on cooperation in the prisoner's dilemma. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, Pennsylvania, April 1969.

(With Margaret G. Hermann) Negotiation in leader and delegate groups. *Journal of Conflict Resolution*, 1968, 12, 332-344.

(With F. T. Morgan) Task and motivational influences on the assessment of creative and intellectual ability in children. Paper presented at the biennial meeting of the Society for Research in Child Development, Santa Monica, California, March 1969.

KRISTOF, WALTER

(With Michael Browne) On the oblique rotation of a factor matrix to a specified pattern. *Psychometrika*, 1969, 34, 237-248.

Estimation of true score and error variance for tests under various equivalence assumptions. Paper presented at the Mathematical Psychology meeting, Stanford, California, August 1968.

*Untersuchungen zur Theorie psychologischen Messens*. Meisenheim (Germany): Hain, 1969.

A general theory of oblique factor rotation. Paper presented at the meeting of the Psychometric Society, Princeton, New Jersey, 1969.

Structural properties and measurement theory of certain sets admitting a concatenation operation. *British Journal of Mathematical and Statistical Psychology*, 1968, 21, 201-229.

LANNHOLM, GERALD V.

(With Gary L. Marco & William B. Schrader) *Cooperative Studies of Predicting Graduate School Success*. Graduate Record Examinations Special Report 68-3, August 1968.

*The Use of GRE Scores and Other Factors in Graduate School Admissions*. Graduate Record Examinations Special Report 68-4, October 1968.

LEAF, WILLIAM A.

(With D. E. Kanouse, J. A. Jones, & R. P. Abelson) Balance, character expression, and the justice principle: An analysis of sentence evaluations. Paper presented at the 76th annual convention of the American Psychological Association, San Francisco, California, 1968.

LEWIS, MICHAEL

(With Marilyn Rausch, Susan Goldberg, & Cornelia Dodd) Error, response time and IQ: Sex differences in cognitive style of preschool children. *Perceptual and Motor Skills*, 1968, 26, 563-568.

(With L. W. Sontag & W. G. Steele) The fetal and maternal cardiac response to environmental stress. *Human Development*, 1969, 12(1), 1-9.

(With Susan Goldberg) The acquisition and violation of expectancy: An experimental paradigm. *Journal of Experimental Psychology*, 1969, 7, 70-80.

(With Susan Goldberg) Perceptual-cognitive development in infancy: A generalized expectancy model as a function of the mother-infant interaction. *Merrill-Palmer Quarterly of Behavior and Development*, 1969, 15(1), 81-100.

Infants' responses to facial stimuli during the first year of life. *Developmental Psychology*, 1969, 1, 75-86.

(With Susan Goldberg) Play behavior in the year-old infant: Early sex differences. *Child Development*, 1969, 40, 21-31.

(With Cornelia Dodd) The cardiac response to a perceptual cognitive task in the young child. Paper presented at the eighth annual meeting of the society for Psychophysiological Research, Washington, D.C., October 1968.

The meaning of an OR: A study in the hierarchical order of attending. Paper presented at the meetings of the Society for Research in Child Development, symposium on The



- orienting response: Issues in developmental inquiry. Santa Monica, California, March 1969; and at the XIXth International Congress of Psychology, London, England, July 1969.
- (With Cornelia Dodd) Attention distribution as a function of complexity and incongruity in the 24-month-old child. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, Pennsylvania, April 1969; and at the Fourth International Seminar on Special Education, Cork, Ireland, September 1969.
- LINN, ROBERT L.  
Computer-assisted testing. Speech presented at the symposium on The computer utility—Implications for higher education, Manchester, New Hampshire, May 5-7, 1969.  
(With T. Anne Cleary & Donald A. Rock) Reproduction of total test score through the use of sequential programmed tests. *Journal of Educational Measurement*, 1968, 5, 183-187.  
(With T. Anne Cleary) A note on the relative sizes of the standard errors of two reliability estimates. *Journal of Educational Measurement*, 1969, 6, 25-28.  
(With T. Anne Cleary) Error of measurement and the power of a statistical test. *British Journal of Mathematical and Statistical Psychology*, 1969, 22, 49-55.  
A note on the stability of the Iowa tests of basic skills. *Journal of Educational Measurement*, 1969, 6, 29-30.  
(With Donald A. Rock & T. Anne Cleary) The development and evaluation of several programmed testing methods. *Educational and Psychological Measurement*, 1969, 29, 129-146.  
(With Charles E. Werts) Analyzing school effects: How to use the same data to support different hypotheses. *American Educational Research Journal*, 1969, 6, 439-447.
- LORD, FREDERIC M.  
Some test theory for tailored testing. Speech presented to the Conference on computer-based learning, instruction, testing, and guidance, Austin, Texas, October 1968.  
Estimating true-score distributions in psychological testing (An empirical Bayes estimation problem). Paper presented at the meeting of the American Association for the Advancement of Science, Dallas, Texas, December 1968.  
An analysis of the Verbal Scholastic Aptitude Test using Birnbaum's three-parameter logistic model. *Educational and Psychological Measurement*, 1968, 28, 989-1020.
- MARCO, GARY L.  
The performance of first-year graduate students on the Graduate Record Examinations. GRE Special Report 68-2, July 1968.  
(With Gerald V. Lannholm & William B. Schrader) Cooperative studies of predicting graduate school success. GRE Special Report 68-3, August 1968.  
A design for evaluating educational programs in a large city. In J. L. Hayman, Jr. (Chm.), Management problems in conducting large-scale research and evaluation studies. Symposium presented at the annual meeting of the American Educational Research Association, Los Angeles, California, February 1969.  
(With Stanley Zdep) A commentary on Kerlinger's structural theory of social attitudes. Research Bulletin 69-49. Princeton, New Jersey: Educational Testing Service, June 1969.
- MEDLEY, DONALD M.  
(With Carolyn Schluck & Nancy Ames) Assessing the learning environment in the classroom: A manual for PROSE recorders. Princeton, New Jersey: Educational Testing Service, December 1968.  
(With R. A. Hill) Goal oriented teaching exercise: Methodology for measuring the effects of teaching strategies. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, California, February 1969.  
OSCAR goes to nursery school: A new technique for observing pupil behavior. Research Memorandum 69-15. Princeton, New Jersey: Educational Testing Service, 1969. (Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, California, February 1969).  
The impact of the classroom. In *Disadvantaged Children and Their First School Experiences, Theoretical Considerations and Measurement Strategies*. Project Report 68-4. Princeton, New Jersey: Educational Testing Service, December 1968.  
(With R. A. Hill) Dimensions of classroom behavior measured by two systems of interaction analysis. *Educational Leadership*, 1969, 26, 821-824.  
The research context and the goals of teacher education. *Educational Comment*. Toledo, Ohio: University of Toledo, 1969.
- MESSICK, SAMUEL  
Response sets. In D. L. Sills (Ed.), *International Encyclopedia of the Social Sciences*. Crowell-Collier, 1968.  
(With L. J. Stricker & D. N. Jackson) Desirability judgments and self-reports as predictors of social behavior. *Journal of Experimental Research in Personality*, 1969, 3, 151-167.  
Measures of cognitive styles and personality and their potential for educational practice. In K. Ingenkamp & T. Marsolek (Eds.), *Developments in Educational Testing*. London: University of London Press, 1968. (Also published in German: In K. Ingenkamp & T. Marsolek (Eds.), *Möglichkeiten und Grenzen der Testanwendung in der Schule*. Berlin: Julius Beltz, 1968.)  
(With D. N. Jackson) A distinction between judgments of frequency and of desirability as determinants of response. *Educational and Psychological Measurement*, 1969, 29, 273-292.  
(With L. J. Stricker & D. N. Jackson) Dimensionality of social influence. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, 189-190.  
(With L. J. Stricker & D. N. Jackson) Evaluating deception in psychological research. *Psychological Bulletin*, 1969, 71, 343-351.  
Can you do real research in the real world? *Untangling the*



- Tangled Web of Education*. Research Memorandum 69-6. Princeton, New Jersey: Educational Testing Service, 1969. (With M. Dunnette (Ed.), B. Baxter, D. W. Bray, A. Carp, R. L. Ebel, and A. G. Wesman) Job testing and the disadvantaged. *American Psychologist*, 1969, 21, 637-650.
- NORTON, DANIEL P.  
The potency of educational treatments. Speech presented to a pre-session of the Wisconsin Education Research Association, December 1967.  
Management emphases in the installation of an evaluation staff for federally-funded projects in a large city. Paper presented at the symposium on Management problems in conducting large-scale research and evaluation studies at the AERA meeting, February 1969.  
Curricular supplements for educational researchers. Speech presented at the American Association of Professors of Educational Research, Summer 1968.  
(With Jayjia Hsia) Evaluation of integration of Evanston District 65 schools—Interim Report, June 1969.
- NOVICK, MELVIN R.  
(With Frederic M. Lord and Allen Birnbaum) *Statistical Theories of Mental Test Scores*. Reading, Massachusetts: Addison-Wesley, 1968.  
Review of J. E. Hill and A. Kerber, Models, methods, and analytical procedures in educational research. *Journal of the Royal Statistical Society, Series A*, 1968, 131, 440-441.  
(With M. A. Woodbury) Maximizing the validity of a test battery as a function of test length for a fixed total testing time. *Journal of Mathematical Psychology*, 1968, 5, 242-259.  
Multiparameter Bayesian indifference procedures (with discussion). *Journal of the Royal Statistical Society, Series B*, 1969, 31, 29-64.  
(With D. T. Thayer) Some applications of procedures for optimally allocating testing time. Research Bulletin 69-1. Princeton, New Jersey: Educational Testing Service, 1969.  
(With D. T. Thayer) A general computer program for optimally allocating testing time. Research Bulletin 69-50. Princeton, New Jersey: Educational Testing Service, 1969.  
A Bayesian approach to the selection of predictor variables. Research Bulletin 69-58. Princeton, New Jersey: Educational Testing Service, 1969.  
(With D. T. Thayer) A study of corrections for restriction of range. Research Memorandum 69-22. Princeton, New Jersey: Educational Testing Service, 1969.
- OWEN, ROGER J.  
A Bayesian analysis of Rasch's multiplicative model for misreadings. Research Bulletin 69-64. Princeton, New Jersey: Educational Testing Service, 1969.  
A Bayesian approach to tailored testing. Research Bulletin 69-92. Princeton, New Jersey: Educational Testing Service, 1969.
- PAYNE, DONALD T.  
(With Mary Ellen Parry & Stephen Harasymiw) Percentage of pupillary dilation as a measure of item difficulty. *Perception & Psychophysics*, 1968, 4 (3), 139-143.
- PENNELL, ROGER  
The influence of communality and N on the sampling distributions of factor loadings. *Psychometrika*, December 1968, 33, 423-440.  
Iterated communalities, convergence, and the number of common factors. *Educational and Psychological Measurement*, Autumn 1968, 28, 793-802.  
True variance in the common factor solution. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1969.  
A new method for estimating common factor scores. Paper presented at the joint meeting of AERA-NCME, Los Angeles, California, February 1969.
- PETERSON, RICHARD E.  
Predictive validity of a brief test of academic aptitude. *Educational and Psychological Measurement*, Summer 1968, 28, 441-444.  
*The scope of organized student protest in 1967-1968*. Princeton, New Jersey: Educational Testing Service, 1968.  
Reform in higher education—Demands of the left and right. *Liberal Education*, March 1969, 55, 60-77. (Speech presented at the annual meeting of the Association of American Colleges, Pittsburgh, Pennsylvania, January 15, 1969)  
(With Bruce K. Eckland) The Exeter study: Technical report and compendium of findings. Progress Report 69-9. Princeton, New Jersey: Educational Testing Service, June 1969.  
Needed instruments for collegiate assessment in the future. Speech presented at the annual meeting of the Educational Research Association of New York State, Concord, New York, November 22, 1968.  
The Institutional Functioning Inventory—Development and uses. Speech presented at the Conference of Liberal Arts College Presidents, sponsored by the Higher Education Center, Temple University, Philadelphia, Pennsylvania, December 3, 1968.  
Organized student protest in 1967-1968. Speech presented at the annual meeting of the American Educational Research Association, February 7, 1969.  
What are college students telling us about alienation? Speech presented at the California Personnel and Guidance Association State Convention, Disneyland, California, February 15, 1969.  
Discussant for a panel on Studies of student life using the College Student Questionnaires. Speech presented at the annual convention of the American Personnel and Guidance Association, Las Vegas, Nevada, March 31, 1969.  
The ETS Study of Boards of Trustees: Some implications. Speech presented at the seventh annual President's Seminar, sponsored by the Junior College Leadership Program, University of California, Berkeley, California, June 22, 1969.
- PIKE, LEWIS W.  
Patterns of effective contextual constraint in sentences. Paper presented at the AERA meeting, February 1969.

ROSENFELD, MICHAEL

(With Harry Harman, Dean Seibel, and Benjamin Shimberg) *Evaluation of driver education and training programs*. Princeton, New Jersey: Educational Testing Service, 1969.

The impact of the school: Measurement strategies. *Disadvantaged children and their first school experiences*. Project Report 68-4. Princeton, New Jersey: Educational Testing Service, 1968.

Negro-white differences in intellectual growth. Paper presented at the meeting of the American Psychological Association, September 1968.

An evaluation of a summer reading institute. Project Report 69-3. Princeton, New Jersey: Educational Testing Service, 1969.

(With Thomas Hilton) Negro-white differences in adolescent educational growth. Research Bulletin 69-71. Princeton, New Jersey: Educational Testing Service, 1969.

SEIBEL, DEAN W.

Measurement of aptitude and achievement. In D. K. Whitla (Ed.), *Handbook of measurement and assessment in behavioral sciences*. Reading, Massachusetts: Addison-Wesley, 1968.

(With Harry Harman, Michael Rosenfeld, and Benjamin Shimberg) *Evaluation of driver education and training programs*. A Report from the Highway Research Board of the Division of Engineering, National Research Council, National Academy of Sciences, National Academy of Engineering to the National Highway Safety Bureau of the Federal Highway Administration, U.S. Department of Transportation, March 31, 1969.

STRICKER, LAWRENCE J.

"Test-wiseness" on personality scales. *Journal of Applied Psychology*, 1969, 53. (Monograph Supplement, No. 3.)

(With Samuel J. Messick and D. N. Jackson) Desirability judgments and self-reports as predictors of social behavior. *Journal of Experimental Research in Personality*, 1968, 3, 151-167.

(With Samuel J. Messick and D. N. Jackson) Evaluating deception in psychological research. *Psychological Bulletin*, 1969, 71, 343-351.

TANAKA, MASAKO N.

(With Anne M. Bussis, Jean H. Orost, and M. Papanek) Teaching classification concepts to disadvantaged children. Paper presented at the meeting of the American Educational Research Association, Los Angeles, California, February 1969.

THOMPSON, RAYMOND E.

Introductory Physical Science Achievement Tests. *Introductory Physical Science, Physical Science II, A Progress Report*. Newton, Massachusetts: Education Development Center, Fall 1968.

A survey of the teaching of physics in secondary schools. Test Development Report 69-3. Princeton, New Jersey: Educational Testing Service, January 1969.

(With Guenter Schwartz) The College Board Physics

Achievement Test. *The Physics Teacher*. March 1969, 7(3).

TRISMEN, DONALD A.

Management and generalizability. Paper presented at the annual meeting of AERA, symposium on Management problems in conducting large-scale research and evaluation studies, February 1969.

WARD, WILLIAM C.

Creativity in young children. *Child Development*, 1968, 39, 737-754.

Reflection-impulsivity in kindergarten children. *Child Development*, 1968, 39, 867-874.

Rate and uniqueness in children's creative responding. Research Bulletin 68-36. Princeton, New Jersey: Educational Testing Service, 1968. (*Child Development*, 1969, 40, 869-878, and as a speech presented at the 1969 American Psychological Association convention)

Creativity and environment cues in nursery school children. Research Bulletin 68-58. Princeton, New Jersey: Educational Testing Service, 1968. (*Developmental Psychology*, 1969, 1, 543-547, and as a paper presented at the 1969 SRCD meetings)

WARREN, JONATHAN R.

The interaction between grading procedures and instructional philosophies. Paper presented at the Conference on Evaluation in Medical Education, Edmonton, Alberta, Canada, June 1969.

WERTS, CHARLES

(With Robert L. Linn) Analyzing school effects: How to use the same data to support different hypotheses. *American Educational Research Journal*, 1969, 6, 439-447.

WILSON, CORNELIA D.

(With Michael Lewis) Attention distribution as a function of complexity and incongruity in the 24-month-old child. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, Pennsylvania, April 1969.

WINGERSKY, BARY G.

Classification problem. Paper presented at the Western Psychological Association Convention, Vancouver, British Columbia, Canada, June 1969.

WOHLWILL, JOACHIM F.

(With J. H. Flavell) Formal and functional aspects of cognitive development. In D. Elkind, & J. H. Flavell (Eds.), *Studies in cognitive development*. Oxford University Press, 1969

Methodology and research strategy in the study of developmental change. In P. B. Baltes and L. R. Goulet (Eds.), *Theory and research in life-span developmental psychology*. New York: Academic Press, in press.

The age variable in psychological research. *Psychological Review*, 1970, 77, 49-64.

Research on development of concrete operations. A col-

loquium, Department of Psychology, Princeton University; also Department of Psychology, City University of New York.

Age: Independent variable, dependent variable, or what?

A colloquium, Department of Psychology, George Washington University, Washington, D.C.

Man as a seeker and neutralizer of stimulation. Paper pre-

sented at the Sanitarists' Institute on Environmental Quality Management, University of Connecticut, Storrs, Connecticut.

Methodology and research strategy in the study of developmental change. Paper presented at the Conference on Life-Span Developmental Psychology, University of West Virginia, Morgantown, West Virginia.